

Treatment of patients with personality disorders

**Day hospital treatment followed by combined psychotherapy
compared with outpatient individual treatment**

Espen Arnevik

**Ullevål Personality Project,
Department for Personality Psychiatry and
Department for Research and Education,
Ullevål University Hospital.
Department of Psychology,
University of Oslo.
Norway**

© **Espen Arnevik, 2009**

*Series of dissertations submitted to the
Faculty of Social Sciences, University of Oslo
No. 178*

ISSN 1504-3991

All rights reserved. No part of this publication may be
reproduced or transmitted, in any form or by any means, without permission.

Cover: Inger Sandved Anfinsen.
Printed in Norway: AiT e-dit AS, Oslo, 2009.

Produced in co-operation with Unipub AS.
The thesis is produced by Unipub AS merely in connection with the
thesis defence. Kindly direct all inquiries regarding the thesis to the copyright
holder or the unit which grants the doctorate.

*Unipub AS is owned by
The University Foundation for Student Life (SiO)*

Table of Contents

| | |
|--|----|
| Summary | 5 |
| List of papers | 9 |
| Abbreviations | 11 |
| 1. Introduction | 13 |
| <i>Table 1: The general diagnostic criteria for a Personality Disorder</i> | 15 |
| 1.1 Treatment of personality disorders | 16 |
| 1.2 Effectiveness of treatment | 17 |
| 1.3 Personality dysfunction and assessment of change | 23 |
| 1.4 Cost-effectiveness of treatment | 25 |
| 1.5 Summary | 28 |
| 2. Aims of the study | 29 |
| 2.1 Paper I: Validation of the SIPP-118 | 29 |
| 2.2 Paper II: Short-term change..... | 30 |
| 2.3 Paper III: Long-term change | 30 |
| 2.4 Paper IV: Health utilization costs..... | 31 |
| 3. Methods | 33 |
| 3.1 Ullevål Personality Project – design | 33 |
| <i>Figure 1: Study design Ullevål Personality Project</i> | 33 |
| 3.2 Therapy..... | 34 |
| Combined psychotherapy (CP) | 34 |
| Out patient individual psychotherapy (OIP) | 36 |
| 3.3 Therapists | 36 |
| Individual therapists | 36 |
| Group therapists | 37 |
| 3.4 Assessments..... | 38 |
| Axis I diagnoses | 38 |
| Axis II diagnoses | 38 |
| Clinical measures | 39 |
| Psychosocial functioning..... | 40 |
| Costs | 41 |
| <i>Table 2: Unit costs</i> | 42 |

| | |
|---|-----|
| 3.5 Procedures | 42 |
| Evaluation..... | 42 |
| Randomization | 43 |
| Ethics | 43 |
| Medication..... | 43 |
| Statistics | 44 |
| 3.6 Subjects | 46 |
| Exclusion criteria..... | 46 |
| <i>Figure 2: Number of patients participating in follow-up evaluation</i> | 47 |
| Completeness of data | 47 |
| Patient description at baseline | 48 |
| <i>Table 3: Sociodemographic and clinical characteristics of the 114 participants at baseline</i> | 49 |
| <i>Figure 3: Axis II diagnoses (%)</i> | 50 |
| <i>Figure 4: Axis I diagnoses (%)</i> | 51 |
| The Dutch samples | 51 |
| 4. Summary of papers..... | 53 |
| 5. Discussion | 57 |
| 5.1 Methodological issues | 57 |
| Participants..... | 57 |
| Treatments | 60 |
| Methods..... | 61 |
| Summary of the methodological strengths and limitations | 68 |
| 5.2 Discussion of the results..... | 68 |
| 6. Conclusions and future directions | 77 |
| 7. References | 79 |
| Attachments..... | 91 |
| Paper I: | 91 |
| Paper II: | 115 |
| Paper III:..... | 135 |
| Paper IV: | 155 |

Summary

Personality disorders (PDs) affect functional impairment, and the personal burden of the disease and costs associated with the disorder are considerable. Although the last years well-designed studies leads to treatment optimism, evidence based knowledge regarding optimal format and length of treatment for the wide range of PDs seen in clinical practice is still lacking. The two formats of most interest with respect to both effectiveness and costs might be outpatient individual psychotherapy and various forms of day hospital treatments. The lack of valid assessments measuring dimensional aspects of core PD pathology has been a limitation in most former efficacy studies for patients with PDs. The Severity Indices for Personal Problems (SIPP-118) is such an instrument. Thus, this thesis has two main aims; 1) evaluating the validity of the Norwegian version of the SIPP-118, and 2) comparing two treatment modalities regarding clinical effect and costs.

The Ullevål Personality Project (UPP) is a randomized controlled study (RCT) comparing day hospital treatment followed by combined psychotherapy (CP) with outpatient individual treatment (OIP) for patients with PDs (N=114). In paper I the data from the Norwegian translation of the SIPP-118 was compared with two Dutch samples and relevant clinical measures. In paper II and III, the short-term and long-term efficacy of CP and OIP was compared through extensive assessment of the patients' symptoms, relational functioning, psychosocial functioning and personality pathology at the 8 and 18 months follow-up. In paper IV, the health utilization costs before and during the two treatments were compared.

The cross-national validity study of the SIPP-118 showed adequate psychometric properties at the facet level and seems promising as a measure of PD core pathology. Future studies should investigate the instruments ability to distinguish between axis I and axis II pathology and its sensitivity to change. The main results from the two efficacy studies were that there were no differences in change between CP and OIP. This challenges the notion that extensive and structured day hospital treatment models are more efficient than outpatient individual treatment for patients with PDs. The analysis of subgroups showed no indication that the initial day hospital treatment was better for the most severely disturbed patients. Finally, there was no difference in total costs for patients in CP and OIP as the higher treatment costs in CP were compensated by a reduction in the use of additional health services. Future studies should comprise larger samples of patients and the research should be supplemented with process studies exploring individual and non-linear changes, as well

as predictors of change.

Acknowledgements

This study is a clinical study, and I would first of all like to thank all of the patients participating in this study. There have been many interviews and a large amount of self-report questionnaires to fill out. Thank you!

I would also like to thank the participating therapists for taking their share of filling out questionnaires and forms, but most of all for your open-mindedness, cooperation and for letting researchers evaluate your work.

Thank you, Theresa Wilberg, for all support and guidance. Your thoughtful examinations of the data and the writing, your keen eye for details without losing sight for the big picture, and your red pen containing criticism, firm structure and touching consideration has been of tremendous help and inspiration. I also owe thanks to all the participants in the research group led by Sigmund Karterud at the Department of personality psychiatry. The ambitious and visionary, yet open-minded and inclusive atmosphere has been inspirational. Additionally, the competence in this group has been inspiring and instructive both regarding clinical work, research and project management. I also wish to thank Jon Monsen for support and guidance throughout this work.

I owe thanks to many people whose support and presence has made these years memorable. Geir Pedersen has been sitting in the office next to me, the door has always been open, and he has always had time for a chat about both personal and statistical issues. Christian, the conversations on the staircase has been significant in many ways! To all my family and friends, who is always there, and Annmari who came into my life making the finishing stage a joy. And most of all I wish to mention my beautiful little girl Kaja, who repeatedly reminds me of the amazing fact that, regardless of my professional aspirations, most of all I am a father.

The financial support from Helse Øst, Josef and Haldis Andresen's Legacy, the Institute of Psychiatry, University of Oslo and the Department of psychiatry, Ullevål University Hospital made this thesis possible.

List of papers

Paper I:

A cross-national validity study of the Severity Indices of Personality Problems (SIPP-118).

Arnevik E, Wilberg T, Andrea H, Monsen J.T, Karterud S.

Ref. type: Personality and Mental health. In press, 2008.

Paper II:

Psychotherapy for personality disorders: Short term day hospital psychotherapy versus outpatient individual therapy - a randomized controlled study.

Arnevik E, Wilberg T, Urnes O, Johansen M, Monsen J.T, Karterud S.

Ref. type: European Psychiatry. In press, 2008.

Paper III:

Psychotherapy for personality disorders: 18 months' follow-up of the Ullevål Personality Project.

Arnevik E, Wilberg T, Urnes O, Johansen M, Monsen J.T, Karterud S.

Ref. type: Submitted Journal of Personality Disorder, 2008.

Paper IV:

Health service costs for patients with personality disorders treated with combined psychotherapy versus individual psychotherapy: 18 months follow-up of the Ullevål Personality Project.

Arnevik E, Halsteinli V, Karterud S, Wilberg T.

Ref. type: Prepared for submitting European Psychiatry.

Abbreviations

| | |
|---------|--|
| AIC | Akaike's information criterion |
| APA | American Psychiatric Association |
| APD | Avoidant personality disorder |
| BDI | Beck depression inventory |
| BHS | Beck hopelessness scale |
| BPD | Borderline personality disorder |
| CBGT | Cognitive behavioural group treatment |
| CBT | Cognitive behavioural treatment |
| CIP | Circumplex of personality problems |
| COI | Cost of illness |
| CP | Combined psychotherapy |
| DAPP-BQ | Dimensional assessment of personality pathology |
| DBT | Dialectic behaviour therapy |
| DHP | Day hospital treatment |
| DSM | Diagnostic and Statistical Manual of Mental Disorders |
| ES | Effect size |
| FFM | Five factor model |
| GAF | Global Assessment of Functioning |
| GAF-F | Global Assessment of Functioning – function score |
| GAF-S | Global Assessment of Functioning – symptom score |
| GLMM | General linear mixed modelling |
| GSi | General Symptom Index (SCL-90-R) |
| ICC | Intra class correlation |
| IIP | Index of interpersonal problems |
| IIP-C | Index of interpersonal problems - Circumplex |
| ISE | Index of self-esteem |
| MBT | Mentalization based treatment |
| MINI | Mini-International Neuropsychiatric Interview for axis I diagnoses |
| NOK | Norwegian kroner |
| OIP | Outpatient individual psychotherapy |
| PD | Personality disorder |

| | |
|----------|---|
| PD NOS | Personality disorder not otherwise specified |
| RCT | Randomized controlled trial |
| SCID-II | Structured Clinical Interview for DSM-IV, axis II |
| SCL-90-R | Revised symptom checklist – 90 item version |
| SD | Standard deviation |
| SFT | Schema focused therapy |
| SIPP-118 | Severity indices of personality problems – 118 items |
| SIPP-SF | Severity indices of personality problems – short form |
| TAU | Treatment as usual |
| TCI | Temperament and character inventory |
| TFP | Transference focused psychotherapy |
| UPP | Ullevål Personality Project |

1. Introduction

Personality disorders (PDs) are complex disorders leading to significant distress and impairment for the individuals affected. There has been a growing literature on treatment relevant theories for PDs, but there is still a lack of knowledge regarding efficacy of various treatments and treatment modalities for PDs, the assessment of PDs, cost-effectiveness of PD treatments, and efficacious factors in the processes towards change. The Ullevål Personality Project (UPP) is a randomized controlled trial (RCT) addressing some of these questions. This thesis is a part of the UPP and focus specifically on the validation of one specific instrument assessing PDs, short-term and long-term efficacy of two modalities of treatment for PDs, and the health utilization costs during these treatments. The results from the UPP could be highly relevant for clinical work with patients with PDs, as well as in health political decision making.

Measuring a problem usually involves measuring incidence and prevalence. The intention beyond these measurements is to define the populations at risk and adopt indicators that correspond to the problem definition. Prevalence estimates for PDs in the community from five studies during the late 1980's and early 1990's range between 6% and 23% (Torgersen, 1995). A later review of community surveys reported a prevalence of 4.4% - 13% (Coid, 2003). Analyses based on the Australian National Survey of Mental Health and Wellbeing, estimated a lifetime prevalence of PD of 6.5% (Jackson & Burgess, 2000). Data from the 43093 participants in the American 2001 to 2002 National Epidemiological Survey on Alcohol and Related Conditions showed PD prevalence between 13.6% and 19.7% (Dawson et al., 2005). Data from Nordic countries are sparse, but in a Swedish sample of 1000 adults from Gotland, a PD prevalence of 11.1 % were reported (Ekselius et al., 2001). In a Norwegian population of 2054 participants ranging from 18-65 years of age a prevalence of 13.4% were estimated (Torgersen et al., 2001). The prevalence estimates of single disorders in the surveys reported above varies between 0% and 7.9%. Cross cultural differences have been pointed out, and one study shows that the prevalence of Avoidant PD (APD) in Norway was two times the usual findings in other studies. Schizoid PD (SPD) and Paranoid PDs (PPD) were also more common, and Borderline PD (BPD) less common than reported from studies in other countries (Torgersen et al., 2001).

For clinical practice and treatment the prevalence of PDs in clinical populations are of importance because of their association with the duration, recurrence, outcome of psychotherapeutic and pharmacological approaches (Bender et al., 2001; Smith & Benjamin, 2002). One large study found that out of 859 psychiatric outpatients 31% (45% including PD NOS) were diagnosed with at least one PD (Zimmerman et al., 2005) while other studies have shown prevalence ranging from 11% to 75% (Melberg et al., 2003; Oldham et al., 1995). According to Gråwe and colleagues, approximately 15% of patients treated by Norwegian private practitioners generally receive a PD diagnosis (Gråwe et al., 2005), probably indicating an under diagnosing of PDs at private practitioners in Norway. There is of course a heterogeneity regarding severity and impairment for persons with PDs. Some individuals with PDs may live well with their disorders while others with more severe impairment may need repeatedly long-term hospitalisation. The most common PDs in psychotherapeutic settings are BPD, APD and NOS, while patients with other PDs such as antisocial PD and schizotypal PD are more often found in prisons or inpatient psychiatric hospitals (Coid et al., 2006). The issue of severity and type of treatment that is most effective for the various PDs is still relatively unexplored. The above mentioned issues will be approached in this thesis, both through the validation of an assessment of severity and in the discussion of efficacy of treatments.

The general diagnostic criteria for a PD are “*an enduring pattern of inner experience and behaviour that deviates markedly from the individual’s culture*” (see table 1). This pattern is manifested in two or more of the following areas: cognition, affectivity, interpersonal functioning or impulse control (American Psychiatric Association, 1994, p. 633). While remembering the large heterogeneity in the PD patient group, PDs may affect the functioning and increase distress in nearly every realm of concern. The disorder affects the patients’ self-experience, cognitive function and social and occupational functioning. In addition, patients with PDs exhibit higher medical health service utilization, are exposed to more violence, and may respond poorer to both somatic and psychiatric treatment. As a group, PD patients also display more symptom disorders as anxiety, mood disorders, suicidality, eating disorders, and drug and alcohol abuse than non-PD patients (Smith & Benjamin, 2002).

Table 1: The general diagnostic criteria for a Personality Disorder (DSM-IV, p. 633)

| |
|---|
| <p>A. An enduring pattern of inner experience and behaviour that deviates markedly from the expectations of the individual's culture. This pattern is manifested in two (or more) of the following areas:</p> <p>(1) Cognition (i.e., ways of perceiving and interpreting self, other people, and events)</p> <p>(2) Affectivity (i.e., the range, intensity, lability, and appropriateness of emotional response)</p> <p>(3) Interpersonal functioning</p> <p>(4) Impulse control</p> <p>B. The enduring pattern is inflexible and pervasive across a broad range of personal and social situations.</p> <p>C. The enduring pattern leads to clinically significant distress or impairment in social, occupational, or other important areas of functioning.</p> <p>D. The pattern is stable and of long duration and its onset can be traced back at least to adolescence or early adulthood.</p> <p>E. The enduring pattern is not better accounted for as manifestation or consequence of another mental disorder.</p> <p>F. The enduring pattern is not due to the direct physiological effects of a substance (e.g., a drug of abuse, a medication) or a general medical condition (e.g., head trauma).</p> |
|---|

There is a common use of the medical term “comorbidity” to describe the co-variation among disorders although our understanding of mental disorder has not yet reached the level where we can describe the disorders as truly distinct. Hence, the recently more used term co-occurrence will be used (Krueger, 2005). Various factors will affect rates of co-occurrence: a) the time frame of the study, b) the sample in question, c) diagnostic methods used, d) criteria overlap, and e) cut-off criteria (Clark et al., 1995). Taken these factors into account, the independence of axis I and axis II diagnoses has been evaluated using odds ratio statistic. As shown in a thorough review by Dolan-Sewell and colleagues (Dolan-Sewell et al., 2001) axis I and axis II disorders tend to co-occur at greater than chance levels. They concluded that “*having an axis II disorder appears to place the patient at risk for an axis I disorder and vice versa*” (p. 99). They found in their review that nearly three

quarters of the patients with an axis II disorder also had a present axis I disorder. The issue of co-occurrence between axis I and axis II disorders is important for the assessment of PDs as well as for clinical work, especially regarding diagnosis specific treatment. This issue will be discussed further in this thesis.

Regarding axis I/axis II co-occurrence an epidemiological survey in a very large sample (N=43093) found a high degree of co-occurrence among the different PD diagnoses (Grant et al., 2005). Diagnoses within each cluster were significantly associated with each other as were some diagnoses across clusters, especially avoidant PD and dependent PD. These co-occurrences may indicate that each of the PD diagnoses may be alternative manifestations of the same underlying disease process, or core pathology.

1.1 Treatment of personality disorders

Personality pathology and the problems associated with it have been recognised long before PDs were formally diagnosed as it is today. Already in the 19th century disorders which today would be described as character pathology or PDs, were addressed by theories and treatments (e.g. Breuer, Freud etc.). The view of personality disorders as inflexible, pervasive, of long duration, and with early onset has been implicit throughout the diagnostic manuals. In DSM-III the personality disorders were separated from symptom disorders in the new multi-axial system. In the revisions made according to research and professional feedback, a set of general diagnostic criteria underscoring the chronic qualities of PDs were introduced (see Table 1: DSM-IV general criteria). A pessimistic view on PDs as chronic disorders resistant to change was supported by clinical research. One early review of 26 studies found in the literature, mainly inpatient treatment, found that only a very small amount of PD patients showed remission over a 15 year period (Perry, 1993). One possible explanation for the common view on PD as a chronic condition is of course the definition in the diagnostic manuals. But methodical difficulties such as deciding PD diagnoses retrospectively based on old clinical journals, and the use of poorly validated measures of change might have influenced the results in a negative way (Perry, 1993). Such a pessimistic view might turn into a self-fulfilling prophecy where neither clinicians nor patients believe in change, and thus no change will happen (Verheul, 2006). When comparing the lack of change found in the early studies with more recent studies showing a natural course of remission in PD pathology (Perry et al., 1999), one might also speculate if

there was an earlier inappropriate treatment practice leading to iatrogenic effects (Fonagy & Bateman, 2006).

In the late 1990's there was a growing clinical optimism as a set of studies showed good effects of treatment. The review by Perry et al. (Perry et al., 1999) concluded that *“psychotherapy is an effective treatment for personality disorders and may be associated with up to a sevenfold faster rate of recovery in comparison with the natural history of disorders”* (p.1312). In addition, Dialectic Behavioural Therapy (DBT) was presented as a new highly structured treatment specifically designed for symptom improvement in BPD patients. In an early RCT they reported that psychotherapy, meaning the DBT model, was an effective treatment for PDs (Linehan et al., 1993). This was followed by other studies showing that symptomatic prognosis was better than previously recognised (Zanarini et al., 2003). In later years, the research on PD treatments has increased tremendously, especially concerning BPD, and the increasingly positive results have led to treatment optimism. On the other hand, most studies have suffered from methodological limitations not allowing firm conclusions from single studies or comparisons across studies (Davidson et al., 2006b). There are numerous theories of PDs based on different concepts and understandings of psychological development, ranging from evolutionary and cognitive concepts, through interpersonal and attachment approaches, to psychoanalytical theories (Lenzenweger & Clarkin, 2005). Even though the development in the field seems to focus on manualized and specialised treatment programs based on some of these specific theories, psychodynamic outpatient individual psychotherapy is still probably the most common treatment for PD patients. Studies in this area are scarce. There is a need for studies addressing the most usually offered PD treatments, taking into account the need for comparison between studies, and at the same time meeting the methodological requirements for studying psychotherapeutic change. The studies reported in this thesis address many of these issues.

1.2 Effectiveness of treatment

The number of psychotherapy studies on PD treatment has increased the last decades. In 1999 Perry and colleagues (Perry et al., 1999) located 15 studies reporting pre-treatment to post-treatment effects, including three randomized controlled trials (RCTs). Their conclusion was that psychotherapy is an effective treatment for PDs and may be associated with a much faster recovery rate in comparison with the natural history of the disorder. Two years later, with an approach comparing cognitive and dynamic therapy, Leichsenring and

Leibing (Leichsenring & Leibing, 2003a) identified 25 studies carried out from 1974 to 2001. They concluded that both cognitive and psychodynamic therapies are effective treatments of PDs. Both reviews point to the limitation in the limited number of studies available, and underlines the need for examining specific forms of psychotherapy for specific types of PDs. Bateman and Fonagy (Bateman & Fonagy, 2000) relies mainly on cohort studies when they conclude that evidence do not suggest superiority of one type of therapy over the other, nor which type or modality of treatment the patients with different PD diagnoses should be offered. They suggest randomized controlled design studies (RCTs) testing long-term, integrated, and theoretically coherent treatment modalities to enhance knowledge in the field.

For the current thesis a literature search of studies published between 1974 and 2007 in Medline, Psycinfo, major textbooks and journal articles was performed. The search revealed 21 RCTs where PD pathology was the main focus of treatment. Most of the studies focused on individual therapy and were concentrated on the treatment of BPD. Today, the most elaborated and structured treatment models accessible for research are Transference focused psychotherapy (TFP), Schema focused treatment (SFT), Mentalization based treatment (MBT), and Dialectic behavioural treatment (DBT). All of these treatments has a primary focus on BPD, although MBT also have shown promising results in one study of patients with BPD and co-occurring antisocial PD (Bateman & Fonagy, 2008b). Cognitive treatment has been most studied as short-term treatment, but one RCT studying long-term cognitive treatment found no additional effect of this treatment compared to treatment as usual (Davidson et al., 2006a).

Two studies have addressed the effectiveness of Transference Focused Psychotherapy (TFP), one which is a comparison with Schema Focused Therapy (SFT) (Giesen-Bloo et al., 2006), and one comparing TFP with DBT and supportive treatment (Clarkin et al., 2004; Clarkin et al., 2007). Both studies show that patients with BPD respond to the dynamic structured treatments in an outpatient setting. Changes were found in multiple domains of outcome, although the drop-out rate was significantly higher in TFP compared to SFT. SFT showed convincingly good results in the Giesen-Bloo study, and although these findings must be replicated, SFT seems promising as a structured treatment for BPD patients. The studies has been criticised for poor external validity as the treatments in question were highly specialised, therapists were trained and supervised by experts throughout the

treatment, and the patients were supposed to benefit from outpatient treatment. Thus, the treatments might be difficult to implement in the health care clinics.

Mentalization based treatment (MBT) has been studied in one RCT comparing a partial hospitalisation program with standard psychiatric care. They found that the superiority of partial hospitalisation was maintained in the 18 months follow-up (Bateman & Fonagy, 2001). At the 8 year follow-up they concluded that longer-term changes were stimulated (Bateman & Fonagy, 2008a).

DBT's specific ability to reduce drop-out, self-injury and need for emergency inpatient treatment has been proven in a number of RCT's (Linehan et al., 1991; Linehan et al., 1993; Linehan et al., 1994; Linehan et al., 2006; Verheul et al., 2003). Although many studies have reviewed the emerging research in the field little is known about predictors of DBT treatment outcome, partly because most of the outcome studies have been with relative small samples that do not provide statistical power for identifying such predictors (Koerner & Linehan, 2000; Lynch et al., 2007; Lynch et al., 2006; Robins & Chapman, 2004).

The above studies of outpatient psychotherapy for BPD, conducted in accordance with specific treatment models and manuals, have all shown significant reductions in dropout from therapy, remission of BPD diagnoses and changes on clinical variables compared to treatment as usual. It may seem as if the selection and consistent application of a coherent and understandable treatment model is likely to be more important than the choice of a specific theoretical treatment reference. On the other hand, all of these studies can be criticised for low generalizability. The treatments are highly specialised and the patient group has mainly comprised BPD patients.

Following the differences in theoretical orientation, focus, and modes of treatment, there is a question of the impact of differences in treatment dosage or length of treatment. The American Psychiatric Association's (APA's) treatment guidelines recommend that BPD patients are given 1) long-term, individually adapted combined treatment in the form of "psychiatric management", treatment with medication, crisis interventions, and psychosocial support, as well as 2) at least one year of psychotherapy (American Psychiatric Association, 2001). Thus, long-term treatment is recommended, although type or modality of treatment is not specified. Taking into account differences in severity of the disorder and type of diagnoses, emerging questions might be how long is long-term treatment, and whether

short-term treatment may be adequate for some patients. So far, these questions are empirically unanswered.

In a recent study with a non-PD sample, patients receiving short-term therapies displayed heightened work ability more quickly than patients in long-term therapy, but long-term therapy seemed more effective than short-term therapies in the long run (Knekt et al., 2008). The unpublished Munich study seems to point in the same direction (Huber et al., unpublished material). As yet, there has not been published any RCTs on short-term versus long-term treatment with PD samples only, but several studies have investigated the effect of short-term treatment for PDs. These studies indicate that both short-term cognitive and short-term dynamic psychotherapy earn its place among the treatment modalities for patients with PDs. The patients in these studies are not as severely disturbed as patients in studies of long-term treatment, and one may speculate if short-term treatment might be especially effective in treating moderately severe patients with PDs in cluster C (Abbass et al., 2008; Emmelkamp et al., 2006; Svartberg et al., 2004).

A dimension adding complexity to the research of PD treatments is the different treatment modalities seen in clinical practice. SFT and TFP are mainly individual approaches, although SFT are also developing for use in psychotherapeutic groups. MBT are used both in groups and individual treatment, and the DBT program consists of parallel group and individual treatment. In addition, patients with PDs receive a variety of inpatient treatments, day hospital treatments, day-care treatments, group treatments and different combinations of these treatment modalities. It has been suggested that more progress may be found in studies that integrate different modes of treatment rather than compare rivaling theoretical orientations (Bartak et al., 2007). For the current thesis, a computerised search for studies of PD treatments was performed (1974 - 2007), and 37 studies were found where other treatment modalities than individual psychotherapy was part of the treatment. An impression based on the latter search was that there is still a lack of evidence-based knowledge regarding the optimal format and length of treatment, and treatments for the mixed spectre of PDs. Verheul and Herbrink (Verheul & Herbrink, 2007) reviewed four different formats and settings for psychotherapy delivery, that is, outpatient group psychotherapy, outpatient individual psychotherapy, inpatient psychotherapy and day hospital psychotherapy. Although they concluded that various psychotherapeutic treatments had proven to be effective in reducing symptoms and personality pathology, the two formats

of most interest with respect to both effectiveness and cost were outpatient individual psychotherapy and various forms of day hospital treatment.

In Norway, various types of day hospital treatment are probably the most developed specialised multimodal treatment for PDs. At present, there are 16 units in the Norwegian Network of Psychiatric Day Hospitals Day (the “Network”) specialised in group based PD treatment (Geir Pedersen, personal communication). Day hospital treatment, a partial hospitalisation in which the patients receive from two to five days of psychotherapy a week but live at home, is a more time-intensive treatment than outpatient individual psychotherapy, but less time intensive than inpatient treatment. It differs from day care in the emphasis given to psychotherapy. Day hospital treatment stands out as a multi-component eclectic treatment that includes cognitive, behavioural, and psychodynamic elements, with a diversity of profiles across hospitals and countries (Kallert et al., 2004a). The use of day hospital treatment has increased in Europe during recent decades (Kallert et al., 2004a). In Germany, there is a trend towards replacing the traditional inpatient units for psychosomatics and psychotherapy with day treatment programmes (Zeech et al., 2005). This trend is supported by a few studies reporting day hospital treatment as being more efficacious and cost-effective than inpatient treatment (Chiesa et al., 2004; Kallert et al., 2004b).

Two randomised controlled studies have reported promising effects of day hospital treatment. Piper and colleagues (Piper et al., 1993), who compared the efficacy of day hospital treatment, lasting four to five months, to a waiting list control group, found a mean difference in effect size of 0.71 in favour of day hospital treatment. Patients in their study sample were diagnosed with PDs or major depression, and showed a moderate level of psychosocial dysfunction. Bateman and colleagues (Bateman & Fonagy, 1999), who studied a sample of more poorly functioning patients with BPD, found that MBT based day hospital treatment with a time limit of 18 months was superior to standard psychiatric care, and that day hospital treatment also showed considerable cost savings after treatment (Bateman, 2003; Bateman & Fonagy, 2008a). Results from naturalistic cohort studies have also shown effects of day hospital treatment for patients with various types of PDs (Karterud et al., 1992; Karterud et al., 2003; Krawitz, 1997; Wilberg et al., 1998; Wilberg et al., 1999). The scarcity of RCTs for day hospital treatment vs. outpatient care for patients with PDs requires further research as pointed out in one Cochrane review (Marshall et al., 2001). They

concluded: “*there is only limited evidence to justify the provision of day treatment programmes and transitional day hospital care, and no evidence to support the provision of day care centres*” (Marshall et al., 2001, p.2). Moreover, there is a need of knowledge concerning the various elements within day hospital treatments. The Ullevål Personality Project (UPP) started in 2004 as a response to the lack of basic research on day hospital programmes pointed out in the Cochrane review.

The most frequently studied treatment programs have been followed by long-term outpatient psychotherapy, which appears to provide good value (Bateman & Fonagy, 2001). Bateman and Fonagy (Bateman & Fonagy, 2001) reported very good outcomes from their outpatient phase consisting of group psychotherapy provided twice a week, while Chiesa and colleagues (Chiesa et al., 2004), who added individual support to their group therapy in the second phase of inpatient treatment, observed good results from this combined treatment approach. On the other hand, in the prospective naturalistic studies of Karterud et al. (Karterud et al., 2003) and Wilberg et al. (Wilberg et al., 2003), short-term day treatment produced good outcomes, whereas only modest gains were obtained from the following outpatient group psychotherapy phase. One explanation of the observed difference in results might be that patients respond to specific modes or doses of treatments according to the severity of their PD diagnoses. More studies are needed to clarify the effects of different modalities, dose and intensity of treatment.

Another limitation in studies of PD treatment is linked to the significantly increased odds of axis I disorders in patients with axis II disorders as found in epidemiological studies (Grant et al., 2004). Studies investigating effectiveness of treatment have mainly used symptom measures for estimating change dimensionally. The most frequently used measures are the symptom check list (SCL-90), Beck depression scale, measures of interpersonal problems, and different forms of other measures correlating high with current symptom distress. Short-term effects probably mainly affect the PD criteria that appear less trait like (McGlashan et al., 2005). Thus, one might speculate if the measures used in most treatment studies are able to detect changes in personality pathology. The use of symptom measures might question the interpretation of the results from these studies. Thus, developing instruments that assess the core pathology of PDs is an important task for further progress in the field.

1.3 Personality dysfunction and assessment of change

The extended use of symptom measures, and lack of measures of personality pathology in PD research and the clinic, has been critically commented throughout the field. One possible explanation of the lack of valid and clinical useful assessments of PD pathology might be the different theories concerning core PD pathology.

One example is the definition of BPD as a disorder in the self-structure brought about through attachment patterns shaped during development (Fonagy et al., 2002). Thus, the disorder is environmentally induced distortions of psychological functioning, which decouples key mental processes necessary for interpersonal and social function. An object relational model would emphasise the degree of differentiation and integration of self and other representations along with affective valence as constituting personality organisation. Kernberg distinguishes between three levels of personality organisation: neurotic, borderline and psychotic (Kernberg, 1984) where most patients with PDs are on the borderline level of organisation. Cognitive theory also has an emphasis on the patient's interpersonal relationships and assumes that the individuals perception and interpretation of situations shapes his or her emotional and behavioural responses to a situation (Beck et al., 1979). Thus, psychopathology is the result of systematic errors, biases and distortions in perceiving and interpreting events, resulting in dysfunctional responses which may have consequences that serve to perpetuate the dysfunctional cognitions (Young, 1999). Instruments assessing the core pathology in accordance with these theories have been developed, or the development is in progress. Such instruments heighten the internal validity of theory specific treatment studies, as they will be able to detect change in the areas where the specialised treatments focus their interventions. On the other hand, the strong internal validity weakens the external validity as such theory specific instruments are not appropriate for comparison between studies with different treatments. Thus, the generalizability of the results would be low. Moreover, regarding clinical utility, most theory specific assessments are based on interviews and therefore time consuming (e.g. the Adult Attachment Interview). In addition, they may require prolonged training for management and scoring. So, there is a challenge to assess PD core pathology cutting across theories and diagnoses and being useful for research as well as in the clinic.

There is a growing acceptance that there is a dimensional continuity between normal personality and personality pathology (Livesley & Jang, 2000; Trull, 2005; Verheul, 2005;

Widiger & Simonsen, 2005). Several researchers hold the view that normal personality comprises constitutionally based temperament, or basic tendencies (McCrae et al., 2000; Skodol et al., 2002b), as well as adaptive capacities developing in interaction with early caregivers and social environment (Livesley & Jang, 2005). In line with this view, the basic tendencies that form personality style are distinguished from the subject's personality function referring to the functioning or malfunctioning of the adaptive capacities. PDs then involve a harmful dysfunction in the normal adaptive functions of the dynamic personality system, and not just statistical deviance on personality trait dimensions. Empirical studies underline this notion as scores on dimensions of normal personality functioning do not appear to be as strongly associated with functional impairment as the psychopathology of DSM personality disorder (Skodol et al., 2005a). It has been suggested that dimensional normal personality traits provide an alternative way to classify personality disorder (Widiger & Costa, Jr., 1994; Widiger & Lowe, 2007). Most of these studies has relied on the five factor model (FFM), which defines a structure characterised by the five higher order traits: extraversion, neuroticism, agreeableness, conscientiousness and openness to experience (McCrae et al., 2000; McCrae et al., 2001). But although studies have demonstrated that the FFM account for a substantial proportion of the variance in PD diagnoses (Nestadt et al., 2008) there are limitations regarding the extent of FFM capturing the complexity of PDs, strength of the correlations between PD diagnoses and FFM traits, and questions are being raised whether the FFM is actually helpful in understanding PDs and in making clinical decisions. Other instruments aiming at a dimensional assessment of personality dysfunction are: the SASB (Benjamin et al., 2006), TCI (Svrakic et al., 2002), DAPP-BQ (Livesley et al., 1998), SWAP-200 (Mullins-Sweatt & Widiger, 2008; Westen & Shedler, 2007), and the SIPP-118 (Verheul et al., 2008), which is the focus of one of the papers in this thesis. Most of the instruments suffer from limitations such as being theory specific, too time consuming because of length of the instrument or because of interview as method, or that the instrument has not been validated in external studies.

At present, there is a lack of valid instruments that measure the common core features of maladaptive personality pathology. Such a measure of "core pathology" should preferably be based on clinical theories of personality development and should focus on common dysfunctions across the existing diagnostic categories. The Severity Indices of Personality Problems (SIPP-118) is a newly developed instrument addressing most of the challenges assessing PD core pathology (Verheul et al., 2008). SIPP-118 is developed in consistency

with the view that the changeability of personality and PDs is likely to be more pronounced for (mal)adaptive capacities than for the more stable constitutionally based characteristics (McGlashan et al., 2005). Adaptive capacities usually refer to the dynamic organisation of personality that concerns the regulation of self and relationships with others, and comprise characteristics like affect- and impulse regulation, self- and other representations, identity, coping strategies, and acquired skills. Thus, a dimensional measure of core personality pathology might be applicable both to clinical work with patients as well as in research on treatment change. As emphasised by Livesley (Livesley & Jang, 2005), adaptive capacities are essential for a subject's ability to fulfil major life tasks. Even if some of the adaptive capacities of personality to some degree may be influenced by biological constitution (McCrae et al., 2000), the conceptual distinction between basic tendencies and adaptive capacities may have heuristic value for the development of a further understanding of the core pathology of PDs. This might be especially important in the understanding of maturation and change in PDs during adult life, and change due to therapeutic interventions.

The SIPP-118 was developed by a research group in the Netherlands in cooperation with an English research group (Andrea et al., 2007). A conceptual model was first developed through consensus meetings of ten clinical experts in the field of personality and PDs. They all agreed with the notion that the concept of changeable core components is more or less synonymous to that of adaptive capacities, with PDs characterised by deficient levels of adaptive capacities. The experts were requested to identify as many specific adaptive capacities as possible and 264 items were originally elaborated. The items were then refined through extensive qualitative evaluation, and quantitative pilot studies (Andrea et al., 2007; Verheul et al., 2008). So far the SIPP-118 has not been tested in PD samples other than those used for the original development of the instrument.

1.4 Cost-effectiveness of treatment

Effectiveness of treatment might not yield all relevant information concerning the treatment in question. Health authorities want to be informed of the cost of treatment. In addition, the knowledge of health service utilization may be clinically relevant in the work with PD patients struggling with interpersonal problems when facing the health services. Various methods to assess costs and benefits exist:

Cost of illness (COI) measures the impact of the economic burden of disease. This is the most comprehensive method for assessing costs, as it estimates the total costs for society

arising as a consequence of the patients' lifestyle, costs of treatment and use of health care services, productivity losses due to morbidity, and productivity costs due to mortality. Such studies can be used to identify avoidable costs or pinpoint the need for development of health promotion, prevention programmes or treatment programmes. In performing a COI, there is a choice between a bottom-up and a top-down approach. The bottom-up approach gathers data from a small sample and extrapolates it to the total population. The top-down approach starts at total cost per healthcare sector and breaks it down to costs of a specific disease. To our knowledge, only one such study has been published in the PD field (van Asselt et al., 2007).

Cost-benefit analysis is a complete evaluation involving all sectors of society. It is ambitious, difficult and rare within health or other public sectors. The resulting information allows resources between different sectors to be evaluated, for example between health and education or between addiction treatment and building a new highway.

Cost-utility analysis attempts to determine deployment of resources within a programme. The costs of alternatives are measured and the consequences are estimated in indexes of health gain. The mostly used index is Quality of life years (QALY) which is based on the estimation of the number of years of life that would be added by a medical intervention. Each year in perfect health is assigned the value of 1.0 down to a value of 0.0 for death. If the extra years would not be lived in full health, for example if the patient would lose a limb, or be blind or be confined to a wheelchair, then the extra life-years are given a value between 0 and 1 to account for this.

Cost-effectiveness analysis compares the effectiveness of different treatment offers by taking into account the cost of the implementation of the treatment and the effectiveness of the intervention.

There has been an increased interest in the field of cost analyses the last years. However, cost studies require careful consideration regarding methodological issues and relevant studies are still scarce. There are one formal COIs (van Asselt et al., 2007), a few cost-utility studies, and various studies which have assessed the health care utilization and cost of alternative treatments. There is still a need for more knowledge in this field.

PDs are associated with poor health-related lifestyle choices, more exposure to violence, a higher risk of suffering from chronic physical conditions, and symptom disorders like

anxiety, mood disorders, suicidality, eating disorders, drug, and alcohol abuse. PDs are also associated with the use of more medication, and patients may show poorer response to both somatic and psychiatric treatment (Frankenburg & Zanarini, 2004; Smith & Benjamin, 2002). In addition, PD patients have more extensive histories of psychiatric outpatient, inpatient, and psychopharmacological treatment than e.g. patients with major depressive disorder and no PDs (Bender et al., 2001; Bender et al., 2006). The impairment following PDs have been well documented (Skodol et al., 2002a; Skodol et al., 2005c; Stolk et al., 2002), and the “burden of disease” in patients with PDs have been compared with the burden in severe somatic illnesses, such as Parkinson’s disease or rheumatic disease (Soeteman et al., 2005). From the perspective of society, the social costs associated with PDs are significant. Rendu (Rendu et al., 2002) found that non-healthcare costs were higher for patients with PDs compared to non-PD patients, although PDs were not independently associated with higher costs as the interaction between PDs and common mental disorders were the significant predictor. The above studies showing both healthcare costs and non-healthcare costs for personality PDs in general and BPD in special, indicates substantial societal costs for this group of patients. This was confirmed in a prevalence-based study of BPD patients from the Netherlands estimating the total societal cost, including all healthcare costs, medication, informal care, productivity losses, and out-of pocket expenses (van Asselt et al., 2007). They found that cost per BPD patient were €16,852 resulting in a yearly cost of illness of €2,222,763,789 where 22% were healthcare related.

As described, there is now evidence that psychotherapy in general is an effective treatment for BPD showing good results in reducing drop-out from therapy, remission of PD diagnoses, and change on clinical variables (Bateman & Fonagy, 2001; Clarkin et al., 2007; Giesen-Bloo et al., 2006; Svartberg et al., 2004). Most of the treatments are resource intensive long-term treatments, but taking into account the substantial societal costs for these patients one can argue that treatment will pay off. One Australian study found that based on the decrease in hospital treatment alone, psychotherapy for BPD patients resulted in a total saving of \$8,431 Australian dollars per patient for one year of treatment (Stevenson, 1999).

Chiesa and colleagues (Chiesa et al., 2002) compared three different health services; one 12 month inpatient hospital program, one “step-down” program consisting of 6 months of inpatient treatment followed by 12 months of outpatient therapy, and one program with

routinely available community services. Significant savings were achieved in the two specialist programs compared to the community services, with the step-down program as the most cost-effective (Beecham et al., 2006; Chiesa et al., 2002). In another study, health care utilization of BPD patients in partial hospital treatment were compared to treatment as usual (TAU) 6 months before treatment, during 18 months of treatment, and over an 18-month follow-up period (Bateman, 2003). The results showed that there were no cost differences between the groups during treatment. The costs of partial hospital treatment were offset by reduced emergency room treatment and less psychiatric inpatient care. Accordingly, during follow-up the partial hospitalisation group showed a decrease in costs which was not apparent in the TAU group (Bateman, 2003). A summary of DBT suggests it might have potential to be cost-effective, although so far the analyses do not support the cost-effectiveness of DBT (Brazier et al., 2006). Palmer and colleagues concluded that although total costs per patient in Cognitive Behaviour Therapy (CBT) were lower than for patients receiving usual care alone, they could not demonstrate any advantage in cost-effectiveness because the CBT group also reported a lower quality of life. In sum, cost-effectiveness studies on PD treatments are still sparse and inconclusive. Limitations regarding the low number of patients in these studies and problems estimating indirect costs make the results only suggestive. Thus, further studies are needed.

1.5 Summary

As a summary we can conclude that there is a high prevalence of PDs in the general population, and a very high prevalence in clinical populations. PDs affect functional impairment to a high degree, and the personal burden of the disease and societal costs associated with the disorders are considerable. Thus, offering effective and cost reducing treatment is in the interest for the individuals, their families and the state. Although there are well-designed studies leading to treatment optimism, the treatment trials still tend to focus on patients with BPD in outpatient psychotherapy. Evidence based knowledge regarding optimal format and length of treatment for the various range of PDs seen in clinical practice is still lacking. The two formats of most interest with respect to both effectiveness and cost might be outpatient individual psychotherapy and various forms of day hospital treatment. Most former treatment studies suffer from the lack of assessing PD pathology. There is a need for the development of instruments measuring dimensional aspects of core PD pathology. SIPP-118 is such an instrument, but has yet to be validated outside the developmental studies.

2. Aims of the study

The Ullevål Personality Project (UPP) aimed at comparing two treatment modalities (CP vs. OIP) for patients with PDs. The CP treatment comprised initial day hospital treatment followed by combined psychotherapy, and the OIP treatment was outpatient individual treatment by private practitioners. The requirement of using not only symptomatic measures but assessing dimensional measures of PD core pathology led to the first main aim of the thesis:

1. Evaluating the validity of the SIPP-118.

Comparing the two treatments implied randomisation of the patients and estimating the efficacy through various clinical measures and health utilization costs. The second main aim of the thesis was:

2. Comparing CP and OIP both regarding clinical effect and costs.

2.1 Paper I: Validation of the SIPP-118

So far the SIPP-118 has not been tested in PD samples other than those used for the original development of the instrument (Verheul et al., 2008). To explore the generalizability of the results from the original Dutch sample, the internal and external validity of the measure needed to be examined. The SIPP-118 was translated to Norwegian and our sample of poorly functioning patients with PDs was compared with two Dutch samples. The aims were to replicate the Dutch study, investigate the SIPP-118 as a valid measure of general personality pathology, and to explore the SIPP-118 as a severity measure for PDs. More specifically, the aims of the present study were:

1. We expected the facets of the SIPP-118 to be internally consistent, and inter-correlations between facets and between domains to be low to moderate.
2. We expected the Norwegian patient sample to be equal to the Dutch PD sample and differ significantly from the Dutch normal sample on SIPP-118 scale scores.
3. We expected a moderate relationship between SIPP-118 and well established clinical measures of psychosocial function and symptomatic distress.

2.2 Paper II: Short-term change

In this first efficacy study of CP and OIP the aim was to examine the short-term effects of the intensive initial treatment in CP with short-term effects in OIP at 8 months follow-up evaluation. We compared the initial 18 weeks day hospital treatment (DHP) to OIP. There are several differences in the treatments being tested. Compared to OIP, the DHP is 1) more intensive, 2) involves more therapeutic elements (e.g. psychodynamic, cognitive, and behavioural therapeutic techniques), and 3) involves a systematic collaboration between several therapists. We expected improvements in both treatments, but due to the differences in the two treatment modalities we expected some increased benefit of the DHP treatment.

1. We expected a lower attrition rate in DHP than in OIP.
2. We expected that fewer patients in DHP would report self-injury and suicide attempts compared to OIP at 8 months follow-up.
3. We expected moderate improvements on clinical measures in both treatments, but with a significant difference in favour of DHP.
4. We expected that patients with more severe personality pathology would benefit more from DHP than from OIP.

2.3 Paper III: Long-term change

At 18 months follow-up we expected positive long-term effects in CP in line with other studies finding small short-term effects but more extensive long-term effects (Bateman & Fonagy, 2008a; Blum et al., 2008). Thus, our general assumption was that the patients would benefit more from the intensive combined treatment following day hospital treatment (CP), than from the OIP. More specifically, the aims of the 18 months follow-up study were:

1. We expected improvement on the wide range of clinical measures with a significant difference in favour of CP.
2. We expected a lower termination rate in CP than in OIP.
3. We also examined if there were specific PD criteria interacting with outcome in the two treatments.

2.4 Paper IV: Health utilization costs

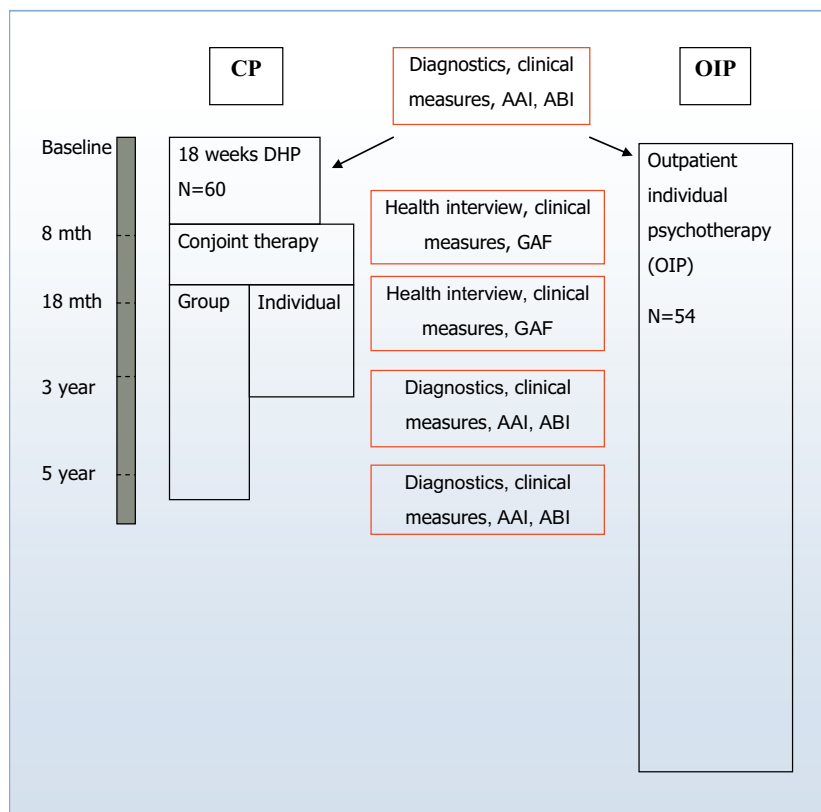
The aim of the study was to compare health utilization costs in CP to health utilization costs in OIP at 8 months follow-up and 18 months follow-up. There is no doubt that the extensive CP treatment is more costly than OIP in direct treatment costs, especially in the initial phase where the patients in CP received intensive day hospital treatment. Yet, comparable studies have shown that higher treatment costs may be compensated by reduced use of additional health care services already during treatment (Bateman, 2003). Hence, the aims were:

1. We expected a reduction in additional health care costs in CP compared to OIP at 8 months follow up, and we expected this reduction to remain at the 18 months follow-up.
2. Thus, we expected the total costs for the CP patients to be similar to OIP at 8 months follow-up.
3. We expected that at 18 months follow-up the total costs in CP would be lower compared to OIP.

3. Methods

3.1 Ullevål Personality Project – design

Figure 1: Study design Ullevål Personality Project



The Ullevål Personality Project (UPP) started in 2004 as an answer to the lack of basic research on day hospital programmes pointed out in the Cochrane review (Marshall et al., 2001). The project is carried out at Ullevål University Hospital, Oslo, Norway by Department for Personality Psychiatry in cooperation with Department for Research and Education and the University of Oslo. The study is a randomised controlled trial (RCT) comparing two modalities of treatment for patients with various PD diagnoses. Consecutively evaluated patients who agreed to participate in the study were randomly allocated to one of the following treatment conditions: a) Combined Psychotherapy (CP) which consisted of initial day hospital treatment followed by long-term combined group and

individual psychotherapy, or b) Outpatient individual psychotherapy (OIP). All patients were evaluated on a wide range of clinical measures at baseline, 8 months, 18 months, 36 months (in progress) and 72 months (starting 2010). All patients were diagnosed on axis I and axis II at baseline, and will be re-diagnosed at the 36 and 72 months evaluations. In addition, an extensive battery of clinical measures and interviews were used. The study design is described in Figure 1.

3.2 Therapy

Combined psychotherapy (CP)

Combined Psychotherapy (CP) is a long-term treatment defined in UPP as the combination of day hospital treatment (DHP) followed by conjoint individual and group psychotherapy.

Day hospital psychotherapy (DHP)

The patients allocated to CP were initially offered 18 weeks day hospital treatment (DHP). The DHP treatment consisted of a combination of psychodynamic and cognitive-behavioural group therapies 3-4 days a week, and was adhering to relational psychodynamics with reference to group analysis, self-psychology and mentalization based treatment. The staff received biweekly video based supervision by a senior therapist. Although written guidelines outlining the therapeutic stance was elaborated, there was no formal manual which could serve as a standard for measuring treatment adherence. The day treatment program had capacity for 18 patients, and all groups had a maximum of nine patients and two therapists. Due to difficult research logistics following the RCT procedures, the actual number of patients was lower for some periods. The treatment program is considered rather typical for the day hospital treatment tradition in Norway (Karterud et al., 2003).

The psychodynamic group therapy was considered to be the core of the treatment program and conducted according to modified group analytic principles. It was provided 1,5 hour twice a week, with an emphasis on maladaptive interpersonal transactions, affect dysregulation, attachment issues, self object needs and self object failures.

The art group therapy was also provided 1,5 hour twice a week and were conducted according to guidelines described by Johns & Karterud (Johns & Karterud, 2004).

The cognitive group therapy was provided one hour once a week. The group worked with early maladaptive schemas which had been identified by the self report instrument Young-75 (Young, 1999).

The problem solving group therapy focused upon strategies for solving social, economic and interpersonal problems and lasted for one hour once a week.

The cognitive-behavioural group therapy (CBGT) for patients who suffered from additional anxiety disorders was optional. It lasted for 1,5 hours once a week and was conducted according to generally accepted CBGT guidelines which included individualised exposure home lessons (Clark, 1999; Hoffart et al., 1993; Marks, 1987).

The median group therapy (“community meetings”) assembled all patients and several staff members for one hour twice a week. It addressed general issues relevant to treatment alliance, i.e. treatment ideology, group norms, adherence to the program, cooperation and the dynamics of the unit as a whole.

Pharmacotherapy: A majority of the patients (70%) were on medication by referral and the medication was evaluated and eventually modified by the staff psychiatrist who also monitored the pharmacotherapy during the treatment. The APA BPD guidelines (American Psychiatric Association, 2001) were followed with a somewhat overall restrictive attitude.

The staff dynamic group met one hour biweekly with an external consultant. Working with patients with PDs is well known to provoke counter transference reactions and disagreements among staff (Rossberg et al., 2007). The primary task of this group was to restore the vitality of the staff and cultivate a sound collective reflective culture.

Staff meetings were held three times a week, orchestrating ongoing events, mutual information and ad hoc supervision.

Follow-up conjoint treatment

After ending the short-term day hospital treatment all patients were offered continuous outpatient treatment. The mean length of the transition phase between day hospital treatment and the following therapy was 19 days (SD = 19). The follow-up treatment offered was once a week outpatient individual psychotherapy (with a predefined maximum length of 2.5 years) conjoint with 1½ hour group psychotherapy once a week (with a predefined

maximum length of 4 years). Conjoint therapy is defined by not having the same therapist in individual and group sessions. Research on long-term efficacy of this treatment modality is scarce (Karterud et al., 2007). The individual therapy adhered to psychodynamic principles while the group therapy adhered to modified group analytic principles. The group therapists and the individual therapists met approximately twice a year to address process and progress issues.

Out patient individual psychotherapy (OIP)

The patients allocated to OIP received eclectic individual therapy defined as therapy according to their usual practice and in agreement between patient and therapist, mainly conducted in private practice. The researchers gave no instructions to the OIP therapists regarding the duration and intensity of psychotherapy, nor did they interfere with any treatment decisions in the OIP condition. Thus, length of treatment was according to consecutive evaluation of the treatment, and the agreement between therapist and patient. The duration and treatment attrition will be discussed later in the thesis. The frequency of therapy ranged from once a month to three times a week, with 83% of the patients attending therapy once a week. This is comparable with the average treatment frequency in outpatient psychotherapy provided by private practitioners in Norway, which has been reported as one consultation a week with a mean treatment duration of 1.5 to 2.0 years (Gråwe et al., 2005; Husum et al., 2005).

3.3 Therapists

Individual therapists

To recruit individual therapists, a mail was conducted to all private practitioners in Oslo who had a contract with the State Health Insurance Fund, as well as professionals working at mental health centres. After a short description of the project, the specialists were asked if they were willing to partake, either as experimental (CP) therapists or as control (OIP) therapists. The therapists who responded were assigned to the experimental or control therapy group according to their own preferences. The treatment expenses for both treatments were covered by the State Health Insurance Fund. At the beginning of the study, the therapists completed a self-report questionnaire covering theoretical preference, education, work experience, and work satisfaction. When the patient(s) terminated treatment, a second questionnaire was administered to obtain reports of frequency of

treatment, duration of treatment, and the reason for ending therapy. There will be a discussion of these issues later in the thesis.

OIP individual therapists

Thirty-two external therapists were recruited as control therapists (16 psychologists, 15 psychiatrists and one resident) who provided psychotherapy consistent with their usual practice style. Each therapist was allocated between one and three patients, and signed a formal treatment contract with the project. The OIP therapists comprised 12 females and 20 males with a mean age of 55 years ($SD = 7.6$). Their mean work experience as psychotherapists was 20 years ($SD = 8.0$). Most therapists reported adherence to psychoanalytic/psychodynamic theories, although cognitive and systemic elements were present. The therapists were generally satisfied with their work as a psychotherapist, as indicated by a mean of 4 ($SD = 0.6$) on a six-point satisfaction scale ranging from 0, no satisfaction, to 5, very satisfied. The therapists were invited to attend an annual one-day motivational seminar on general aspects of PDs.

CP individual therapists

The individual therapists involved in the follow-up outpatient CP treatment comprised 16 psychologists, 12 psychiatrists, 2 psychiatric nurses and one social worker (24 external therapists and seven therapists from the Department of Personality Psychiatry). The mean age of the individual therapists was 50 years ($SD = 8$), and 58% were female. Each therapist received one to three patients for treatment. The therapists were generally very satisfied with their work as psychotherapists, indicated by a mean of 4.2 ($SD = 0.5$) on a six-point satisfaction scale ranging from 0, no satisfaction, to 5, very satisfied. Their mean work experience as psychotherapists was 16 years ($SD = 8.0$). Although there was written guidelines outlining the therapeutic stance adhering to self-psychology and mentalization, no formal adherence was tested, and most therapists probably delivered therapy according to their usual practice. The therapists were invited to one-day seminars twice a year on conjoint therapy, and they were invited to receive monthly supervision in groups, focusing on conjoint treatment.

Group therapists

The group therapists both in the day hospital treatment and the following conjoint therapy were regular staff from the Department of Personality Psychiatry, Ullevål University Hospital. In the day hospital treatment there were 5.5 full-time positions divided among ten

group therapists. There were three psychiatric nurses, two psychiatrists, a residential doctor, an art therapist, a physiotherapist, a social worker and a psychologist. Seven of the ten therapists were trained group analysts (five-year training). The therapists' mean age was 48 years ($SD = 9$), and 80% were female. These therapists also conducted the follow-up outpatient group therapies.

3.4 Assessments

Axis I diagnoses

Symptom disorder diagnoses were based on the Mini International Neuropsychiatric Interview (M.I.N.I) for axis I diagnoses (Sheehan et al., 1994), and decided according to DSM-IV (American Psychiatric Association, 1994). Twenty-five videotaped interviews were rated by an independent rater. The kappa values for seven axis I diagnoses, represented by at least five cases, were .51 for major depression, .60 for dysthymic disorder, .92 for social phobia, .52 for obsessive-compulsive disorder, .51 for panic with agoraphobia, .41 for general anxiety disorder and .52 for alcohol misuse disorders. Altogether, the 25 patients received 88 axis I diagnoses, and there was agreement with respect to 57% of these diagnoses. When pooling the diagnoses, the kappa values were .58 for the patients having any anxiety disorder and .71 for any substance misuse disorder. There was agreement for 23 of the 25 patients who had any type of mood disorder, but the kappa value could not be computed due to empty cells in the cross-tabulation.

Axis II diagnoses

PD diagnoses were decided according to DSM-IV criteria using the SCID-II interview (First, 1994). Twenty-four videotaped SCID-II interviews were rated by an independent rater. The kappa values for three PDs, represented by at least five cases, were .75 for avoidant PD, .66 for borderline PD, and .71 for paranoid PD, indicating acceptable diagnostic reliability. Altogether, the 24 patients received 37 diagnoses of specific PDs. There was agreement with respect to 68% of the diagnoses. The number of fulfilled SCID-II criteria were used as a dimensional severity measure (Skodol et al., 2005b). The intraclass correlation coefficient (ICC 3.2) on number of fulfilled SCID-II criteria was .95 (95% CI: 0.73 – 1.00).

Clinical measures

Self-injury, suicidal thoughts and suicide attempts

Self-injury and suicide attempts were assessed by the patients self-report, and then quality checked using research interviews and revised if necessary. The criteria for self-injury were an intended episode of cutting, burning, etc. without suicidal intent, whereas a suicide attempt was defined as an episode with the intention to commit suicide, but the patient failed to do so. Suicidal thoughts during the previous seven days were reported as being present, or not, and their severity was measured on a 5 point scale with a range from “...passing thoughts” to “...did active preparations for suicide”.

Symptom distress

The *symptom checklist, SCL-90-R* (Derogatis, 1983), was used to measure symptom distress. SCL-90-R, a self-report questionnaire requiring responses on a 0–4 Likert scale, was designed to cover the major symptoms of psychiatric distress, represented by nine dimensions that can be meaningfully expressed by a Global Severity Index (GSI). A higher score indicates more symptomatic distress. The *Beck Depression Inventory (BDI)* (Beck et al., 1961; Groth-Marnat, 1990) was used to measure characteristic attitudes and symptoms of depression. The BDI is a 21-question self-report inventory, with a 0–3 rating scale. A sum score (0–63) is obtained by adding up the scores for each of the questions. Sum scores between 19 and 29 are regarded as indicating moderate to severe depression, while scores of 30 and above indicate severe depression. The *Beck Hopelessness scale (BHS)* (Beck & Steer, 1988) is a self-report instrument containing 20 true–false statements designed to assess the extent of positive and negative beliefs about the future during the past week. Each of the 20 statements is scored as either 0 or 1. A total score is expressed counting the pessimistic responses for each of the 20 items. The total score ranges from 0 to 20, higher scores indicating a greater level of hope. For the assessment of self-esteem, we applied the *Index of Self-Esteem (ISE)* which assesses the self evaluative aspects of self-esteem (Hudson, 1982; Walmyr Publishing CO, 1992). ISE is a 25-item questionnaire measuring the degree or severity of a subject’s self-esteem problems. The scale produces scores ranging from 0 to 100, with scores above 30 indicating clinical significant problems in this area. A 10-point scale assessed subjective *Quality of life*, a score of 1 representing the least perceived quality of life and a score of 10 indicating the highest perceived quality of life.

Interpersonal problems

Self-reported interpersonal problems were assessed by the *Circumplex of Interpersonal Problems (CIP)* (Pedersen, 2000). The CIP is a 0–4 Likert scale, 47-item version of the Inventory of Interpersonal Problems (IIP) (Horowitz et al., 1988). A total mean score (CIP) was generated to summarise all scores on the subscales. A higher score indicated more severe interpersonal problems. The total CIP score correlates .99 with the sum score obtained from Alden's IIP-C (Alden et al., 1990; Pedersen, 2000).

Personality problems

A preliminary exploratory version of the *Severity Indices of Personality Problems (SIPP)* consisting of 203 items was translated from English to Norwegian by our research group, and then back translated into English by a bilingual translator. The final Norwegian translation was determined by group consensus after comparison with the English back translated version. In the present studies, we used 118 items similar to the final SIPP-118 version, with a 0–4 point response scale being used (Verheul et al., 2008). The original study showed that the measure comprised 16 facets that were fairly internally consistent and uni-dimensional. These facets are clustered into five higher order domains named Social concordance, Self-control, Identity integration, Responsibility and Relational functioning, which are weighted sums using primary and secondary loadings in accordance with factor analyses and qualitative considerations (Andrea et al., 2007). This means that the 16 facets contribute to different higher domains with different weights. A description of the higher order domains and the facets with the highest loadings on each domain are shown in paper I. The SIPP-118 domains were employed as measures of (mal)adaptive personality functioning. High scores on these domains indicate higher adaptive functioning, whereas low scores indicate more deficient levels of adaptive capacity. The validity of this measure of core PD pathology is part of this thesis, and will be thoroughly discussed.

Psychosocial functioning

Global Assessment of Functioning (GAF) was applied to assess psychosocial functioning. GAF is rated on a scale from 0 to 100, a higher score indicates a higher level of functioning. GAF was rated according to a split version, that is, symptoms (GAF-S) and function (GAF-F) were assessed separately (Goldman et al., 1992). According to DSM-IV, the GAF score should be determined with respect to the lowest level of functioning. Thus, only the lower of the two scores was used in the analyses. At baseline, GAF was rated by the staff member

responsible for the patient's evaluation after obtaining anamnestic information, completing diagnostic interviews (i.e., SCID-II and MINI), and complemented by a 15-minute GAF interview developed for the present project (<http://personlighetsprosjekt.com/gaf/>). A consensus was then attained at the staff meeting. At 8 and 18 months follow-up, the GAF score was rated by the evaluator accompanied by a consensus method consulting a senior GAF expert. For reliability purposes, all GAF interviews at follow-up were recorded and scored by two independent evaluators gaining consensus. The reliability for GAF measured by Intra Class Correlation (ICC, 3.2) was .94 at 8 months, and .93 at 18 months evaluation.

GAF-F can be compared to Social and Occupational functioning assessment scale (SOFAS) (American Psychiatric Association, 1994) (s.760). GAF-F score 60 was used as a clinical relevant cut off score. The GAF scoring manual states that GAF scores above 60 indicates: "...some difficulty in social, occupational or school functioning, but generally functioning pretty well...", while GAF scores under 60 is defined by a "...moderate difficulty in social, occupational, or school functioning...". To reassure clinical significance of the GAF-F scores, reliable change based on Jacobson and Truax' criteria was added as cut-off criteria (Jacobson & Truax, 1991). Based on the reliability of the instrument (Pedersen et al., 2007a), and the standard deviation in our data, a change in GAF-F score of 5 or more was estimated as reliable change. Reliability for GAF-F with cut-off score 60 was $\Phi = .87$, which indicates a high degree of consensus between the interviewers and the independent scorers in distinguishing between GAF-F level below 60 and GAF-F level above 60.

Costs

Cost related data were collected from each patient using a health service interview adapted for this population and research context. The interview assessed services that were likely to bear the major costs (see Table 2 for unit costs). Data were collected for a retrospective period of 12 months at baseline, the last 8 months at 8 months follow-up, and the last 10 months at 18 months follow-up. For both groups, this gave a continuous picture of the use of health services throughout the study period. Mean cost per day at day hospital treatment was calculated from data on number of full time equivalent therapists and administrative staff, registration of personnel allocation of time, and annual accounts from the Department of Personality Psychiatry at Ullevål University hospital. Cost per private outpatient consultation was estimated from information on annual activity and income from participating therapists. Other units costs were obtained from published reports (Halsteinli et

al., 2003; NLWA, 2008; Pedersen, 2007b; Petersen, 2007) except community day centre, psychiatric nurse and social service visit which were estimated specifically for this project on the basis of information from one large municipality. Medication costs were calculated by means of standardised price information from the Norwegian Medicines Agency. All unit costs are from 2006 when the project started. All costs are presented in Euro, the exchange rate used were 1 Euro = 8 NOK.

Table 2: Unit costs

| | Unit costs (euro) |
|--|----------------------|
| Treatment costs | |
| Day hospital - per day | 165 |
| Outpatient public - per consultation individual | 118 |
| Outpatient public - per consultation group | 59 |
| Outpatient private - per consultation individual | 110 |
| Outpatient private - per consultation group | 45 |
| Additional health care costs | |
| General practitioner - per visit | 33 |
| Psychiatric emergency - per visit | 163 |
| Other emergency services - per visit | 39 |
| Psychiatric hospital - per inpatient 24h stay ¹ | 891 |
| Community mental health centre - per inpatient 24h stay ¹ | 619 |
| Community day centre per visit | 43 |
| Psychiatric nurse - per home visit | 73 |
| Social service - per visit | 28 |

1) Day visit 1/2 of 24h stay

3.5 Procedures

Evaluation

Initial diagnostic evaluation and assessments were performed by the regular staff at the Department of Personality Psychiatry, Ullevål University Hospital. Initial evaluation lasted

between 3-5 sessions for each patient. Final decisions regarding diagnoses and inclusion in the project was done at a clinical meeting between the evaluator and one of two senior psychiatrists at the department. The 8 months evaluation was carried out by Espen Arnevik, and the 18 months evaluation was carried out by a research assistant studying for a master in psychology.

Randomization

The assurance of the randomization schedule generated by SPSS was maintained by a randomization coordinator who was not involved in the patient evaluations. After the initial evaluation of the patients and their agreement to participate in the study, the evaluator contacted the randomization coordinator to receive a randomization number and information regarding treatment allocation. If the patient was randomized to OIP, the evaluator contacted an individual therapist so that outpatient psychotherapy could be initiated. The patients allocated to CP were placed on a waiting list for day hospital treatment. Three patients expressed dissatisfaction with the allocation procedure and withdrew after randomization, one from the CP condition and two from the OIP condition. Thus, the total number of patients included in the baseline sample was 114. Sixty patients were randomly allocated to the CP group, and 54 to the OIP group.

Ethics

The patients in the CP treatment were presumed to receive an extensive and good treatment offer. Despite the differences in treatment dosage, the patients in OIP also received an offer of comprehensive treatment. On average, the patients received treatment that was more comprehensive than “treatment as usual”, especially by Norwegian standards. In Oslo, poorly functioning patients with PDs may have difficulty gaining access to regular psychotherapy within private specialist practices. All participation was based on written informed consent. The project is approved by The Data Inspectorate and Regional Ethics Committee. No procedures in the project were considered potentially harmful. Publication of the study results was not in any way contingent on the sponsor’s approval or censorship of the manuscript.

Medication

Psychopharmacological consultations with a psychiatrist were part of the initial evaluations for all attending patients. In the DHP, medication was monitored during the treatment, and modified by the staff psychiatrist when necessary. The APA BPD guidelines (American

Psychiatric Association, 2001) were followed with a somewhat overall restrictive attitude. In the following CP outpatient treatment collaboration was established with general practitioners regarding medication, medical certificates, occupational rehabilitation, etc. The OIP therapists were instructed to cooperate with general practitioners and other health services according to their usual practice. At follow-up evaluations all patients were offered a consultation with a psychiatrist at the Department of Personality Psychiatry.

Statistics

All statistics were calculated using SPSS 15.0, SPSS 16.0 or STATA /IC 10.0 for Windows.

Pearson's correlation coefficient was applied to test associations between continuous variables, and Spearman's rank-order correlation coefficient was used to test associations between dichotomous variables and continuous variables that were not normally distributed. Chronbach's alpha was used to measure inter item consistency of the SIPP-118 facets.

Five measures were used to test the external validity of the SIPP-118; GAF, CIP, SCL-90-R, total SCID-II criteria, and BDI. Because the number of measures increased the risk of type I error, a Bonferroni correction was calculated based on the number of external measures. The calculations increased the significance level from 0.05 to 0.01. Standard multiple regression was conducted with each global measure as a dependent variable, and the SIPP-118 domains that showed statistically significant correlations with the dependent variable at the $p < 0.05$ level in bivariate analysis. Similar analyses were conducted for the CIP and SCL-90-R subscales as dependent variables. The level of statistical significance was 0.05 for all regression analysis. To compare the SIPP-118 facet profiles in the patient samples from Norway and the Netherlands, and to measure differences between the sample profiles with different PD diagnoses, all facets were included simultaneously. Because the number of facets can be considered to be different measures, a two-group between-subjects multivariate analysis of variance (MANOVA) was used.

For statistical power analysis in the comparison of OIP and CP, we used methods recommended for repeated measures design (Faul et al., 2007). With alpha level of 0.05 and a mean correlation among repeated measures of 0.35, we had a 95% power to detect an effect size of 0.26 reflecting the difference between the two treatments. The study applied an intention to treat approach.

Pre-post effect sizes were computed by Cohen's d , using the pooled standard deviation and adjusting for sample size (Cohen, 1988).

Reliable change was estimated to enhance the clinical relevance of the results comparing the two treatments. Reliable change is an estimate of change based on the reliability of the instrument and the standard deviation of the data (Jacobson & Truax, 1991). Because of the lack of test-retest reliability in our samples, the test-retest reliability estimates were based on the following studies: GSI = 0.86 (Groth-Marnat, 1990), CIP = 0.92 (Pedersen et al., 2007b), GAF = 0.82 (Pedersen et al., 2007a), Quality of life = 0.79 (Pedersen, 2007a), Beck Hopelessness Scale = 0.85 (Holden & Fekken, 1988), BDI = 0.82 (Moffett & Radenhausen, 1990), and SIPP-118 = 0.92–0.97 (Andrea et al., 2007). Differences in the number of patients showing a reliable positive or negative change in clinical measures in the two treatments, compared with patients who did not show any change, were tested by analysing 2x2 frequency tables using Chi-square tests.

In order to strengthen the validity of the efficacy studies, inferences were based on broader patterns of change, rather than single tests. To estimate differences between the treatment conditions at 8 months follow-up the sum scores from the clinical measures were entered into two-group between-subjects multivariate analysis of variance (MANOVA). In addition to this over all test, univariate ANOVA were used to estimate differences on each measure. Multiple logistic regression analyses were used to predict the occurrence of suicidal thoughts and self-injury at 8 months follow-up. The predictor variable was treatment condition (DHP and OIP), with suicidal thoughts or self-injury at baseline included as covariates to adjust for pre-treatment levels of these variables.

To compare efficacy for subgroups of patients at 8 months follow-up, a selection of the most poorly functioning patients in the sample was extracted based on a combination of their capacity for work and the total number of fulfilled SCID II criteria at baseline. This subgroup yielded 26 patients (DHP: $N = 11$, OIP: $N = 15$). MANOVA statistics usually require about $N = 20$ in each of the four cells. Hence, independent sample t -tests for each of the clinical measures were used to test the differences in outcome between treatments for this subgroup of patients.

For the comparison of the two treatments at 18 months follow-up, we used a General Linear Mixed Model (SPSS 16.0). The main advantage of the GLMM approach over standard

repeated multivariate analysis of variance is that it allows for inclusion of cases with missing values, thereby providing a better estimate of the true (unbiased) effect within the intention to treat sample. To examine the efficacy of the treatments, we used a model with time (three time points), treatments (CP and OIP), and time x treatment interaction. Each of the 13 clinical measures was used as dependent variables. The number of fulfilled SCID II criteria, and capacity for work or study prior to inclusion, were added as covariates to estimate the effect of initial differences in severity at baseline. To assess possible confounding impact of premature termination of treatment, termination before 18 months evaluation were also added as covariates. We used Compound Symmetry Heterogeneous as covariate type, as we expected same amount of dependability in covariates, and change in variations across time. Akaikes' information criterion (AIC) was used to inspect the fit of the model. To explore if the number of specific PD criteria interacted with outcome in the two treatments we used the same GLMM approach with a model with time, treatment, and treatment x number of specific PD criteria interaction. The total number of fulfilled SCID II criteria was used as covariate.

In the cost analysis the distribution of continuous variables was examined by comparing mean and normal probability plots (Q-Q plots). Bootstrapped confidence intervals based on 2000 replicates were calculated due to highly skewed cost data, using the statistical package of STATA /IC 10.0 for Windows. The basic idea of bootstrap involves repeated random sampling with replacement from the original data producing random samples of the same size as the original sample. Each of these bootstrap samples provides an estimate of the mean. Repeating the process 2000 times, as in our paper, provides information of the variability of the estimator. A 95% confidence interval for the mean would then be values from 2.5% to 97.5% of this variability.

3.6 Subjects

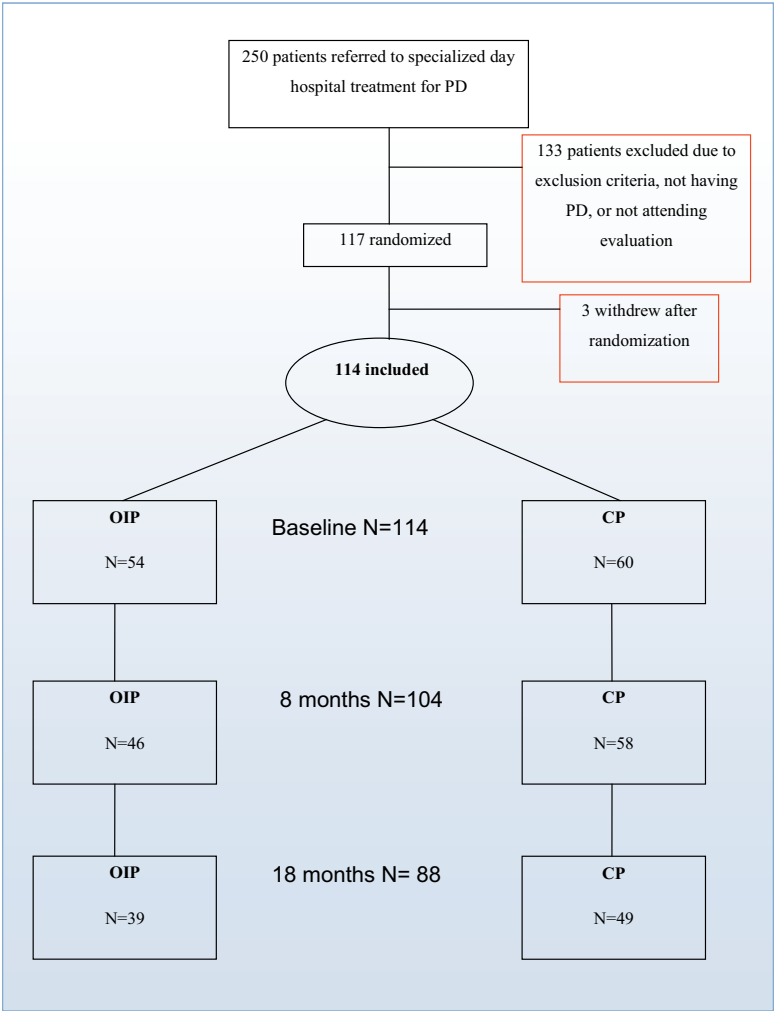
Two hundred and fifty patients were referred to the Department of Personality Psychiatry during the intake period for the study (see Figure 2). Only patients with PDs were included in the study.

Exclusion criteria

Exclusion criteria were schizotypal PD, antisocial PD, ongoing alcohol or drug dependence, psychotic disorder, bipolar I disorder, untreated ADHD (adult type), developmental disorder (e.g. Asperger's syndrome), organic syndromes, and being homeless. Of the 250 referred

patients, 117 met the entry criteria for the UPP. Three patients withdrew after the randomization procedure, and the total number of patients in the sample was 114.

Figure 2: Number of patients participating in follow-up evaluation



Completeness of data

Of the 114 patients included in the sample, 104 (91%) attended the 8 months evaluation and 88 (77%) attended the 18 months evaluation (see Figure 2). Eighty-three patients (73%) completed all three evaluations while 5% (4 CP patients and 2 OIP patients) completed only baseline evaluation. Four of these patients were diagnosed with APD at baseline, two of which had co-occurring BPD, one patient was diagnosed with PD NOS, and one with

obsessive compulsive PD. Five out of these six patients were female. As a group they obviously had more avoidant PD (APD) criteria than patients attending the additional 8 and/or 18 months evaluations (mean = 4.3 vs. 2.5, $F=4.53$, $p<0.05$), they exhibited less dependent PD criteria (mean = 0.5 vs. 1.8, $F=4.65$, $p<0.05$), a lower experienced quality of life (mean = 2.5 vs. 3.5, $F=4.61$, $p<0.05$) and a higher score on interpersonal problems (CIP) (mean = 1.9 vs. 1.7, $F=4.95$, $p<0.05$). In addition, there was a non-significant trend on the various clinical measures suggesting that these patients perceived more severe impairment than patients attending evaluations. When comparing patients attending 18 months evaluation ($N=88$) and patients not attending 18 months evaluation ($N=26$) no such trend was found. There was only one statistically significant mean difference between these two groups. The patients not attending 18 months follow up had a higher CIP score (mean = 1.9 vs. 1.6, $F= 7.8$, $p<0.05$). There were not found any statistical significant differences in reliable change on clinical variables at 8 months follow-up between the groups attending vs. not attending the 18 months follow-up.

Patient description at baseline

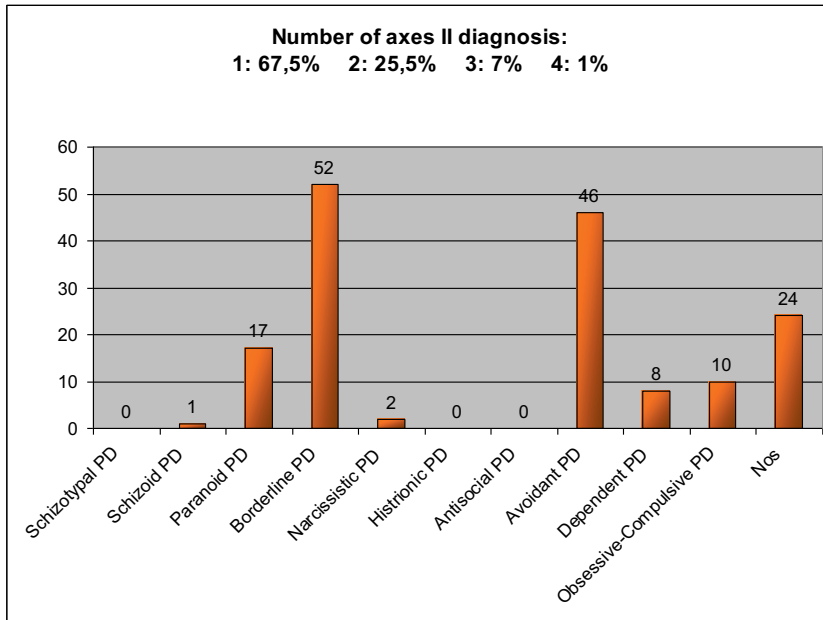
A complete baseline description of the sample is shown in Table 3. The 114 included patients had a mean age of 31 years ($SD = 7.4$), with the oldest patient being 54 years, and the youngest being 20 years of age. A high proportion of the sample was female (74%). Fifty percent of the patients had less than 12 years of education which is lower than average in Norway (<http://www.ssb.no/ssp/utg/200605/05/tab-2006-12-11-01.html>). About 38% of the patients lived alone, and 6% was married. Half of the patients were unemployed, and 21% received support by the social welfare system. Very few defined themselves as home wives, or worked part time. Thirteen percent of the patients reported suicidal attempts during the last 12 months and 31 % reported self harming activities in the same period. Eighty-six percent reported suicidal thoughts during the last 12 months, while 50% reported suicidal thoughts during the last seven days. Mean GSI score was 1.7 ($SD = 0.7$), mean CIP score was 1.7 ($SD = 0.5$), mean BDI score was 19.1 ($SD = 8.8$), mean quality of life score was 3.4 ($SD = 1.6$) and mean GAF score was 47.6 ($SD = 4.4$). Seventy percent of the patients were using, or had been using psychotropic medication last 12 months.

Table 3: Sociodemographic and clinical characteristics of the 114 participants at baseline

| | Total (N=114) | CP (N=60) | OIP (N=54) |
|--|----------------------|------------------|-------------------|
| Age, Mean (SD) | 31 (7.4) | 31 (7.3) | 31 (7.5) |
| Women | 74 % | 77 % | 72 % |
| More than 12 years of education | 50 % | 41 % | 57 % |
| Marital status | | | |
| Live alone | 38 % | 39 % | 38 % |
| Live alone with children | 12 % | 10 % | 13 % |
| Live with parents | 6 % | 3 % | 9 % |
| Common law partner | 12 % | 12 % | 11 % |
| Married | 6 % | 5 % | 8 % |
| Other | 26 % | 31 % | 21 % |
| Work status | | | |
| Full time employee (100%) | 23 % | 22 % | 24 % |
| Half time employee (aprox. 50%) | 8 % | 12 % | 4 % |
| Student | 16 % | 16 % | 16 % |
| Unemployed | 50 % | 50 % | 50 % |
| Other | 3 % | 0 % | 6 % |
| Self injury/suicide attempts | | | |
| Suicidal thoughts last 7 days | 50 % | 51 % | 49 % |
| Suicidal thoughts last 12 months | 86 % | 88 % | 83 % |
| Self harming activities last 12 months | 31 % | 36 % | 26 % |
| Self harming activities lifetime | 48 % | 50 % | 46 % |

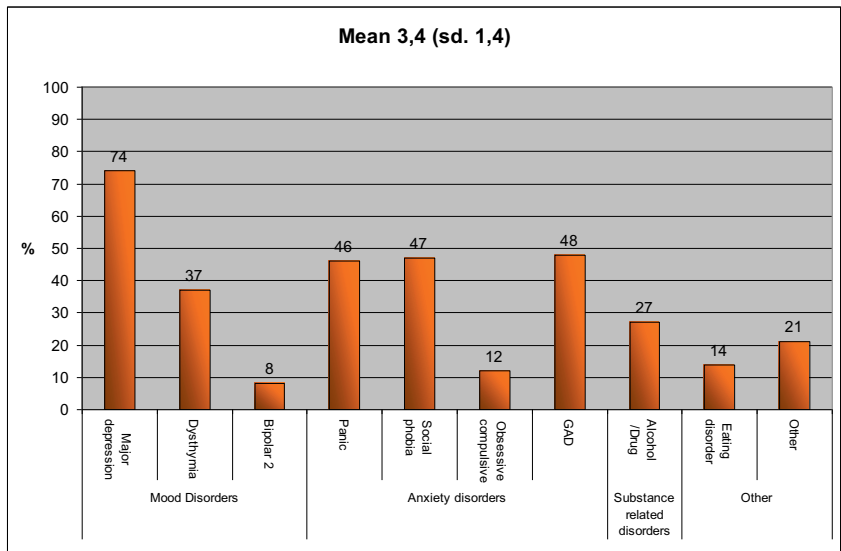
Sixty-eight percent of the patients were diagnosed with one PD, 25% with two, and 7% had three or more co-occurring PD diagnoses. The distribution of the PD diagnoses (see Figure 3) was as follows: borderline PD (46%), avoidant PD (40%), PD not otherwise specified (21%), paranoid PD (15%), obsessive compulsive PD (9%), dependent PD (7%), narcissistic PD (2%), and schizoid PD (1%). The mean number of SCID II criteria was 15.8 (SD = 6.4).

Figure 3: Axis II diagnoses (%)



The mean number of symptom disorders in the sample was 3.4 (SD = 1.4). Seventy four percent of the patients were diagnosed with Major depression, 37% with Dysthymia, 8% with Bipolar II disorder, 46% with Panic disorder, 47% with Social phobia, 12% with Obsessive Compulsive disorder, 48% with GAD, 27% with Substance misuse disorders, and 14% with Eating disorders (see Figure 4).

Figure 4: Axis I diagnoses (%)



The data indicates a sample of patients with moderate to high level of impairment. There were no significant differences between the patients in the two treatment conditions at baseline concerning sociodemographic characteristics, clinical variables, or distribution of PD diagnoses and symptom disorders.

The Dutch samples

In the validation study of the SIPP-118 we used two Dutch samples from the initial studies of the SIPP-118 for comparison (Andrea et al., 2007). The first sample consisted of 478 individuals from the general population, who participated in a postal personality survey. In total, 1520 general community subjects (50% females) from four age groups (15-24; 25-34; 35-44; and 45-54 years) were randomly drawn from the patient files of 15 general practitioners from the Dutch cities and villages Amsterdam (735,500 inhabitants), The Hague (457,700), Tilburg (197,400), Groningen (175,600), Leiden (117,200), Heerlen (95,000), Kerkrade (50,700), Waddinxveen (26,900), Ermelo (26,800), Reusel (12,400), and Laren (11,900). These subjects received by mail several self-report questionnaires including the Severity Indices of Personality Problems (SIPP). Taking into account 35 booklets that were undeliverable, the response rate was 32.2%. Respondents were mostly female (67.6%), and had a mean age of 36.0 years with a standard deviation of 11.6. Educational level was low in 19.3%, intermediate in 49.1%, and high in 31.6%.

The second sample (N=555) comprised patients from six mental health care institutes in the Netherlands (Center of Psychotherapy De Viersprong, Halsteren; Altrecht, Utrecht; Zaans Medisch Centrum De Heel, Zaandam; Center of Psychotherapy De Gelderse Roos, Lunteren; GGZWNB, Bergen op Zoom; Center of Psychotherapy Mentrum, Amsterdam). These institutes offer outpatient, day hospital and/or inpatient psychotherapy for patients with personality pathology and/or personality disorders. As part of the standard intake procedure in these institutions, all admissions underwent a routinely distributed assessment battery including self-report questionnaires and a semi-structured interview to measure psychopathology, personality, functional impairments, and treatment history. Of these 555 patients, 60.0% were female. The mean age was 33.9 years (SD=10.4, range 16-66). Educational level was low in 13.0%, intermediate in 59.5% and high in 27.5%.

4. Summary of papers

4.1 Paper I: A cross-national validity study of the Severity Indices of Personality Problems (SIPP-118)

The objective in this study was to test the validity of a new dimensional measure of maladaptive core pathology for Personality Disorders (PDs), the SIPP-118, by comparing a Norwegian sample of 114 patients with PDs with two Dutch samples. In addition, SIPP-118 scores for patients with avoidant PD and borderline PD were compared, and the relationship between scores on the SIPP-118 and commonly used clinical measures were investigated. The results showed good psychometric properties of the SIPP-118 at the facet level. The Norwegian PD sample had scores equal to the Dutch PD sample and significantly below the general population sample. Correlations with other clinical measures were in the low to moderate range. An instrument measuring the core pathology of PDs has been requested in the area of research and clinical work with patients with PDs, and the results of this study supports a good cross-national validity of the SIPP-118. The instrument seems promising as a dimensional instrument for measuring core personality pathology, but the structure of the higher-order domains of the SIPP-118 should be confirmed in larger patient populations and the facet Respect should be reconsidered. Further studies should focus on the ability to differentiate between axis I and axis II populations. Moreover, the instrument's ability as a measure of therapeutic change is not yet established and should be examined further.

4.2 Paper II: Psychotherapy for Personality Disorders: Short-term day hospital psychotherapy versus outpatient individual therapy - a randomized controlled study

This paper presents the results of the 8 month follow-up investigation from the UPP. The initial day hospital psychotherapy (DHP) was compared with outpatient individual psychotherapy (OIP) for patients with personality disorders (N = 114). The main outcome measures were attrition rate, suicide attempts, suicidal thoughts, self-injury, psychosocial functioning, symptom distress, and interpersonal and personality problems. The study showed a low dropout rate and a moderate improvement on a broad range of clinical measures for both treatments. However, there was no indication of the superiority of one treatment over the other. Neither was there any indication that day hospital treatment was better for the most poorly functioning patients. Further studies will follow this group of

patients for the next five years, the results of which may have implications for resource allocation and the organisation of mental health services for patients with personality disorders. To our knowledge, this study is the first RCT comparing short-term day hospital psychotherapy and individual psychotherapy in private practice for a mixed group of PD patients with a severe to moderate degree of impairment. Overall, the treatment conditions were assumed representative of day hospital treatment in Norway, as well as private psychiatry and psychology practice.

4.3 Paper III: Psychotherapy for Personality Disorders: 18 month evaluation in the Ullevål Personality Project – a randomized controlled trial

The paper presents the results of the 18 month follow-up investigation in the UPP. A step-down model (CP) consisting of initial short-term day hospital treatment followed by conjoint group and individual outpatient treatment was compared with outpatient individual psychotherapy (OIP). A mixed group of patients with PDs (N = 114) with severe to moderate degree of impairment was randomized to the two treatment modalities. The patients were evaluated at baseline, 8 months and 18 months on a wide range of clinical measures assessing symptoms, interpersonal problems, psychosocial functioning and personality pathology. The main finding at 18 months' follow-up was that the intensive, multimodal combined psychotherapy (CP) was not more effective than outpatient individual psychotherapy (OIP). On the contrary, there was a trend of better results in the OIP treatment. The study has to be supplemented with a cost-utility analysis before any consideration of implications for health care planning. However, as for efficacy, the study indicates that eclectic psychotherapy provided by private practitioners has an equal or better effect upon patients with PDs compared to the more comprehensive day hospital and outpatient follow-up treatment.

4.4 Paper IV: Health service utilization costs for patients with PDs treated with combined psychotherapy versus individual psychotherapy – 18 months follow-up in the Ullevål Personality Project

Follow-up studies from the UPP have not revealed major differences in effectiveness between the long-term treatment consisting of initial short-term day hospital followed by conjoint group and individual treatment (CP) compared to individual outpatient treatment (OIP). Thus, costs analyses of the treatments are important for both clinical and political

decision-making concerning patients with personality disorders (PDs). This paper examines the costs of treatment, additional health related services, and medication for patients in the two treatment conditions at 8 and 18 months follow-up (N=114). The main finding was that the higher treatment costs for initial day hospital treatment in CP were compensated by a reduction in costs in the use of additional health services. The total costs for patients in CP were equal to the total costs for patients in OIP at both 8 months follow-up and 18 months follow-up. The results emphasise the need for taking additional health care utilization costs into account when making treatment decisions for patients with PDs.

5. Discussion

The importance of an instrument assessing core PD pathology is emphasised in both research and the clinic. The study of the newly developed SIPP-118 is the first validity study outside the original developmental studies. The results may have significance for establishing a focus on the dimensionality of PD pathology in the clinic as well as strengthening the internal validity in efficacy studies in general, and in the UPP specifically. Thus, strengths and limitations of the SIPP-118 study needs to be addressed.

The UPP is to our knowledge the first RCT comparing initial short-time day hospital psychotherapy followed by long-term conjoint therapy (CP) with individual psychotherapy in private practice (OIP) for a mixed group of patients with PDs. The main findings from the present studies were that the extensive combined psychotherapy (CP) was not more effective than outpatient individual psychotherapy (OIP). These results were present for all main clinical measures, including the newly developed measure of core personality, the SIPP-118, attrition rate and health utilization costs during treatment. The results may have implications for resource allocation and the organisation of mental health services for patients with PDs. Thus, both the internal validity and the external validity of the study need to be discussed. Internal validity refers to the degree of difference in content of the two treatments, to what degree the results in these studies can be attributed to the treatments in question, and to what degree the measures are able to measure clinically important aspects of the patients change. External validity refers to in what degree the results can be generalized to other samples and settings.

5.1 Methodological issues

Participants

One of the strengths in the UPP was the thorough diagnostic procedures at baseline, and the various PD diagnoses in the sample. When evaluating the SIPP-118, the variation in PD diagnoses increased the possibility to evaluate the instrument as a measure of core pathology cutting across the different diagnostic categories. A limitation in the validation study of the SIPP-118 was the lack of a Norwegian normal population. Although the Netherlands may not be culturally very dissimilar to Norway, further studies of the SIPP-118 should compare different Norwegian samples. In addition to a Norwegian normal

sample, a population with only axis I symptom disorders as well as patients with supposedly more severe PD disorders should be included.

The randomization of the patients in the UPP was regarded as successful. The patients in the two treatments were comparable at baseline with respect to the distribution of sex, sociodemographic variables, symptoms, interpersonal and personality problems, psychosocial functioning and diagnoses.

The overall completion of data was high considering the patient population (91% at 8 months, and 77% at 18 months), but characterisations of the patients not attending follow-up evaluations might still have affected the results. It is interesting that the six patients only attending baseline evaluation exhibited more avoidant PD criteria, less dependent PD criteria, higher scores on interpersonal problems (CIP), and a lower experienced quality of life than patients attending the additional 8 and/or 18 months evaluations. This patient profile indicates a challenge with regard to forming a therapeutic alliance. Moreover, the higher rate of patients in CP attending follow-up evaluations compared to OIP patients might point to an advantage for the CP treatment where a therapeutic bond was already initiated during the intake interviews as these were located at the Department of Personality Psychiatry. This could have influenced the short-term results in favour of DHP, although this was not evident in the results at the 8 month follow-up. The difference in attending follow-up evaluations may also be explained by the CP patients forming a relationship with the institution as a whole throughout the initial day hospital treatment, thus feeling more responsibility for meeting to evaluations. On the other hand, the missing patients might have deteriorated, or might have experienced lesser change than the patients attending follow-up evaluations. The GLMM approach does not fully account for possible attrition bias, although meeting for evaluation was added as a covariate in the analyses. As such, the difference in attendance at follow-up evaluations may conceal differences in efficacy which, despite the intention to treat design, is regarded as a limitation of these studies.

The patient sample was not restricted to a specific personality type, but contained different PDs, mostly patients with BPD, APD and paranoid PD with various co-occurring PDs, in addition to PD NOS. This is considered a strength, as other studies comprise patients with mainly BPD. On the other hand, the patient sample did not include the entire spectre of PDs. The lack of patients in Cluster A limits the generalisation of the results, but the PDs in our study are being fairly representative of “typical” PD patients who is present at mental health

centres in small and large cities, and in rural areas (Narud et al., 2005). The study sample contained more patients with BPD than in the Norwegian Network of Psychiatric Day Hospitals (the “Network”) as a whole (46% versus 23%, $N = 2248$, Geir Pedersen, personal communication). However, the study sample was not more dysfunctional on clinical measures at admission than patients in the “Network”.

The PD patients in the study had a severe to moderate level of impairment. Thus, they were not as severely disturbed as the patients in the Bateman and Fonagy study (Bateman & Fonagy, 1999). In their study of an 18 months day hospital treatment program, the patients had a mean GSI of 2.5 ($SD = 0.58$) at baseline evaluation as opposed to our sample where the mean GSI at baseline were 1.7. The patients in the UPP were more similar to the patients in the Giesen-Bloo study (Giesen-Bloo et al., 2006) and the Davidson study (Davidson et al., 2006b) both regarding symptom load and level of psychosocial functioning. In these two studies the patients received outpatient individual treatment. If severity is an indicator for the choice of treatment modality for PD patients, our results with the lack of difference in efficacy between treatments may indicate a sample of patients with a level of impairment where some patients might respond to more intensive and structured treatment while others might respond to less extensive treatment. Hence, a mix of PD diagnoses and severity for patients included in the study might neutralise a possible difference between the two treatments. On the other hand, the argument of better effect from the more structured treatment for the most severe patients, should lead to a difference when comparing subgroups. The most poorly functioning patients would be expected to show more improvement in CP while the better functioning patients would be expected to show more improvement in OIP. That was not the case in our study.

Most studies on PDs are specialised treatments for BPD patients. Arguments for the “general” type of CP treatment evaluated in the UPP, in contrast to specific treatment programmes for self-destructive borderline patients, have been that the “general” CP model can be applied “everywhere”, while specialised treatment programmes directed at specific PDs require a larger catchment area. Although the patient sample in the UPP was relatively large compared with other RCTs in this field, the number of patients was low when analysing subgroups (as with severity and differences between diagnostic groups) and variables with low frequency (as with self-destructive behaviour).

Treatments

It is important to emphasise that the aim of the UPP was not to compare specific treatment techniques. Rather, the outcome studies compared modes or formats of treatment that corresponded to psychotherapy as it has been generally practiced in Norway, both in day hospitals and in private practice. Thus, an important question is how representative these conditions were. Compared with the 12 units in the Norwegian Network of Psychiatric Day Hospitals, the DHP was representative in most of its important characteristics (e.g., short-term format of 18 weeks, group therapies only, number of treatment hours a week, general ideology). A common practice for other day hospitals would be to offer continuation treatment using a group psychotherapy format only, which would be accepted by approximately 50% of the patients. The present study offered conjoint group and individual treatment as continuation treatment, with the expectation that all patients would attend such treatment. During the study, the staff complained that they were “not allowed” to be as confrontational as they would otherwise have been, in order to avoid dropouts and to promote the development of attachment. This was one of the elements which were introduced in the written guidelines. In addition, problems with the research logistics meant that the groups were rather small at times and, as a consequence, even more vulnerable to the effects of poor attendance, which affected the group dynamics. The staff felt obliged to proceed with patients who otherwise would have been regarded as poor candidates for continuation group treatment based on evaluations made at the end of the day hospital treatment. The accepted practice assumed an attachment to “the institution as a whole”, which tried to compensate for possible separation trauma after 18 weeks initial day hospital treatment by ensuring that patients would meet a staff person whom they knew in the following outpatient conjoint treatment (but to whom they were not necessarily attached) as the continuation group therapist. Overall, the experiences of the DHP staff indicated that they were a little more frustrated than usual, they felt somewhat less competent, the completion rate was perhaps artificially high, and the group dynamics were somewhat more difficult. Since group cultures and dynamics fluctuate in day hospitals, and there are some “natural” ups and downs, the culture that was studied was possibly in the lower to middle range of an average day hospital programme.

In OIP, all psychiatrists and clinical psychologists ($N = 32$) worked in private practice and had a contract with the State Health Insurance Fund, which implies that patients pay a maximum of approximately €200 per year, irrespective of how much therapy they receive.

There is no restriction on the number of years spent in treatment. Most OIP therapists in the study proved to be experienced and well qualified. After patient randomization, they received a comprehensive evaluation report of the patients they were to treat, which contained anamnestic information, diagnoses, evaluation of personality functioning and test results. Some therapists reported that the initial phase of treatment was influenced by the patient's dissatisfaction with being allocated to the OIP. In general, the OIP therapists expressed satisfaction with the project as a whole, and welcomed this opportunity to partake in a scientific study. They felt the project to be a good opportunity to demonstrate that private practice is a sound and efficient alternative to more costly treatments. Our impression was that the private practitioners were highly motivated, and possibly in the upper range with respect to experience and qualifications. Their "psychotherapy as usual" approach, in this case, implied that they exhibited more than the usual tolerance with respect to cancelled sessions and payment problems. In summary, it is reason to believe that their services were in the upper range of what could be considered "psychotherapy as usual" in the community.

Methods

Design

One of the main strengths of the UPP was that the efficacy studies reported was completed according to the Cochrane review's recommendations of RCTs and intention to treat design (Marshall et al., 2001). The RCT procedure should ensure similarity in the two patient samples on most variables, which was also found when testing the samples at critical baseline characteristics. RCTs have been criticised for low external validity, that is, the risk of studying conditions that are not representative of clinical practice (Seligman, 1995). However, our RCT comparing modes or formats of treatment that were assumed to be fairly representative of day hospital treatment in Norway, as well as private psychiatry and psychology practice, supports generalizability, and meet modern scientific standards.

The study can be criticised for its lack of a no-treatment group, so that, strictly speaking, the design cannot attribute change to the treatment conditions. The observed change in both conditions may reflect factors such as "a natural recovery process" (Perry et al., 1999; Shea et al., 2002), or a regression towards the mean effect. However, for this patient population with a GAF level of less than 50, a "no-treatment condition" is not realistic. In modern Western health care systems, and particularly in the Scandinavian type of welfare state, a

host of other treatment and care facilities could have been called upon for assistance. The real question is whether a “treatment as usual” (TAU) group should have been included. Based on previous knowledge of the effect of psychotherapy for patients with PDs (American Psychiatric Association, 2001; Bateman & Fonagy, 2001; Leichsenring & Leibing, 2003b; Perry et al., 1999), allocating patients at random to a TAU group after five to eight hours of assessments, which also included assessment of motivation, was regarded as being unethical. The ethically sound and realistic alternative to the initial day hospital treatment, which was regarded as “day hospital treatment as usual”, was “psychotherapy as usual” in Oslo. This was a treatment option which existed “out there”, but which these patients did not often consider because of personal and systemic difficulties.

The logistics when starting up the project was challenging. Two hundred and fifty patients were referred to the Department of Personality Psychiatry, mostly from mental health centres, and were allocated to one of the treatment conditions successively. By 8 months evaluation all patients in OIP had been referred to a therapist and the mean duration of treatment was 4.5 months. Several patients had to wait several months before starting treatment. The delay was due to a number of factors; some patients were referred just before summer holidays, were hospitalised before starting treatment, did not meet, or the therapist did not have the possibility to start treatment right away. The same logistic challenge occurred in CP which resulted in few patients in the DHP group treatment at times. If the waiting period before initial DHP lasted more than 3 months, the patient was asked to fill out the self-report questionnaires again, as to control for possible changes during this period. These are conditions which are difficult to control for, and should be taken into consideration when interpreting the results.

One might question the choice of time for the 8 months follow-up evaluation. A 6 months follow-up would have been beneficial in comparing results with e.g. the Bateman study (Bateman & Fonagy, 1999). On the other hand, taking into account the complicated logistics (e.g. to be sure that all DHP patients had ended short-term treatment, and that all OIP patients had been in treatment a reasonable amount of time before follow-up evaluation), 8 months follow-up was decided.

After the initial day hospital treatment, patients in CP were confronted with the challenges of both separation and forming new attachment relationships as there was a transition to the conjoint outpatient treatment. For most CP patients the 8 months evaluation took place

during or just after this transition phase. A previous study indicated that this period, with all of its dynamic implications, is a difficult and vulnerable phase (Hummelen et al., 2007). Thus, it might be a period where the CP patients experienced increasing symptoms. The OIP patients had the benefit of continuing their treatment with the same therapist.

One argument for the use of short-term treatment has been a reduction in the risk of patient regression, and it can be speculated whether the short-term format revives affect and relational problems without allowing enough time for the patient to work constructively on these problems. Instead, while a stable relationship with the therapist or group would be essential for working through the arising relational problems, the transition phase may be experienced as the breaking of therapeutic bonds, and would not be beneficial. On the other hand, a heightened awareness of relational problems for patients who have adopted strategies for avoiding relational challenges might be necessary for working through these problems.

Assessments

The instrument package was extensive comprising both symptoms measures, measures of interpersonal functioning, measures of psychosocial functioning, and measures of personality pathology. In addition, both self-report and interviews were used as methods. In total, these assessments should be beneficial for detecting changes and possible differences between the treatments.

On the other hand, unreliable measures could lead to concluding that relations between treatment and effect does not exist when they actually does (Type II error). With respect to the studies in the thesis, the reliability of axis I and axis II diagnoses is of importance. Since the introduction of DSM-III in 1980, much research effort has been put in valid classification of personality disorders, and a tremendous amount of literature has been written on treatment according to these categories. Studies have shown that individuals may fluctuate around a general level of personality pathology leading to movements below or above a specific threshold over time. Thus, despite the clinical utility, existing diagnostic categories may not in fact be tremendously helpful to clinicians (Livesley, 2001; Verheul, 2005). Many investigators have noted that a categorical classification system may not be optimal for diagnosing personality disorders, and that PDs is better conceptualised dimensionally. However, there is still considerable debate about which dimensional system is the best option, or if different systems can be hierarchically organised within the same

diagnostic manual (Widiger & Simonsen, 2005). As for today, DSM-IV PDs are mostly conceptualised by reflecting combinations of traits rather than the extreme of a single personality trait dimension. As such, with a threshold of five criteria for BPD, there might not be any significant clinical difference between a patient fulfilling any four BPD criteria and a patient fulfilling five. This is of course unfortunate with respect to group comparisons based on diagnostic status. To increase the validity and reliability of the different diagnoses, there was a strong focus on the assessment process in UPP where the evaluators followed specific procedures. Strictly, patients do not have a disorder, they are found to meet the criteria defined in the diagnostic system and judged to warrant the disorder.

The reliability of axis II diagnoses at baseline evaluation was acceptable, although in the low range compared to other studies (Zanarini & Frankenburg, 2001). This was probably partly due to information bias. The evaluator gathered anamnestic information, had a written referral, and received information from carrying out the MINI interview in addition to the SCID II interview, while the independent rater scored only from the videotaped SCID II interview. To judge whether the patient meets each of the criteria is not easily assessed, and the medium axis II reliability may also be due to the potential confusing effects of axis I states as the majority of the patients included in the study were diagnosed with a co-occurring axis I disorder (e.g. 74% were diagnosed with major depressive disorder). The axis I reliability was low, which might also point to the difficulty in differential diagnostics. The low to medium reliability of the diagnoses could have affected the results in the comparison between diagnostic categories in the SIPP-118 paper as the conclusion of SIPP-118 as a valid instrument for assessing core pathology is based on the reliable difference of APD and BPD diagnoses in the sample.

There is a lack of valid measures for assessing severity and core personality pathology in the PD field. Such a measure of “core pathology” should preferably be based on clinical theory of personality development, and focus on impaired adaptive capacities and common dysfunctions cutting across the existing diagnostic categories. Furthermore, it should be a dimensional measure sensitive to treatment changes. Obviously, a self-report format is less labour consuming and would enhance its clinical utility. SIPP-118 was such an instrument, although it had not been validated outside the initial developing studies, and there was no Norwegian version available. The challenge with the external validation of the SIPP-118, as with other attempts to make dimensional measures of severity, is the lack of consensus on

how to validate severity. In our study we compared clinical samples with a normal sample; we used a variety of the most commonly used clinical measures for relational-, symptoms- and general psychosocial functioning. In addition, we used the number of fulfilled SCID-II criteria as a validation measure. The low correlation with the number of fulfilled SCID-II criteria may have a few possible explanations. Firstly, the number of SCID-II criteria may be a poor measure of severity as the SCID-II criteria differ widely in content. Some criteria points to life threatening events such as suicide attempts and drug abuse whereas other criteria is defined by lack of close friends or being easily influenced by others. As such, one could debate the unidimensionality of this scale as a measure of severity. In addition, one might question the number of SCID-II criteria as a valid measure in our study, but the high reliability scores referred is contrary to such a conclusion. On the other hand, the restricted range of the SCID-II criteria found in this sample may affect the size of the correlations. Finally, the low correlation may indicate that the SIPP-118 is a poor measure for assessing severity.

Another way of using diagnoses to address PD severity has been to regard certain PD categories or clusters as more severe than others, summing up the number of PD diagnoses, or a mixture of these two (Skodol et al., 2005a; Tyrer, 2005). Unfortunately we did not have a wider distribution of diagnoses, as we then additionally could have compared SIPP-118 with these attempts to address severity. These issues should be elaborated in further explorations of the instrument.

Statistics

According to Shadish, Cook and Campbell (Shadish et al., 2002) “*statistical conclusion validity concerns two related statistical inferences that affect the co-variation of causal inferences: 1) whether the presumed cause and effect co-vary, and 2) how strongly they co-vary*” (p.42). As to the first of these inferences, we can conclude that the relation exist in the sample when it actually don’t (type I error). On the other hand, we can conclude that it doesn’t exist when it actually does (type II error). Regarding the second inference we can overestimate or underestimate the magnitude of the relation, as well as the degree of confidence in our results. There are several threats to statistical conclusion validity, some of which have already been addressed, such as reliability of measures, missing data and regression to the mean effects. In addition, low statistical power and restriction of range can

lead to type II error, and high statistical power or extraneous variables can lead to type I error as well as overestimation of magnitude.

With respect to the studies in this thesis some concerns are to be noted. Although the study sample was larger than in most comparable studies, the number of patients restricted the statistical methods used. This was especially evident in the SIPP-118 paper where factor analysis would have been useful for reconfirming the facets and higher order domains found in the original studies of the instrument. Generally, the medium size sample could increase the possibility for type II error in all the papers. Power sensitivity analysis recommended for repeated measures design was applied for the efficacy studies (Faul et al., 2007). The estimation showed a 95% power to detect an effect size of .24-.26 reflecting the difference between the two treatments. From this follows that small differences between treatments would not reach significance. In our results there was a trend towards better efficacy in OIP. With a larger sample this trend might have shown statistical significance, although the clinical significance would probably remain low. Of more concern for statistical validity was the restricted range seen in some variables due to the inclusion criteria for the study. The highly skewed data in occurrence of self-injury and suicide attempts, symptoms which are often seen as an essential part of focus for PD treatments, contributed to an uncertainty in efficacy regarding these variables. In addition, the limited range in the total of fulfilled SCID-II criteria might have decreased its validity as an external measure for severity in paper I with the lack of correlation with the SIPP-118 as a result. Similarly, as most of the statistical procedures used in the studies are dependent of the variance in the data, the restricted range in most of the symptom variables, GAF and PD criteria increases the possibility for type II error. Thus, the external validity gained by the sample in this study might reduce the statistical conclusion validity.

The skewness of the health utilization data yields important information and is clinically interesting, but it also affects statistical procedures and interpretation. Some patients reported costs due to use of health services which was very high, up to €200.000 the year before inclusion, while others reported costs as low as €200. Long-term inpatient stays at somatic hospital due to e.g. heart transplantation was the main reason for the high costs for some patients. Therefore, somatic hospital was excluded from the analysis although one may argue that somatic illness and PD may be related, especially if patients were admitted to somatic hospital because of suicide attempts or intoxication. Even though costs related to

somatic hospital was excluded, there were still some high cost services, and high cost patients, which may have influenced the results when comparing the two treatments. Further studies should include larger samples to make subsample analyses possible.

Because of the skewness of the health utilization data we used the bootstrap procedure to produce means and confidence intervals. One main assumption in the bootstrap procedure is that the missing values in the dataset are completely at random. This is a strong assumption and unlikely to hold for the data in our study. On the other hand, one advantage of bootstrap compared to traditional transformation of data is that we can interpret on the original scale and that in transforming data only approximate normal distributions will appear (Cantoni & Ronchetti, 2004). In addition, it has been argued that the relatively simple assumptions required for estimating a valid bootstrapped mean is to be preferred in the analysis of health cost data (Walters & Campbell, 2004).

As for the efficacy studies, a common assumption in psychotherapy research have been that change is gradual and linear. The research designs and statistics used to study change often reflect this assumption. There has been criticism against this assumption and also the fact that the hypothesised predictors of change are measured once or twice and then compared between groups or correlated with symptom change at the end of treatment (Hayes et al., 2007). Most research focuses on group averages, with much less emphasis on the rich information available in individual time course data. Traditional pre-post designs provide only snapshots of the change process, and traditional correlation and ANOVA-based analyses are limited in the extent to which they can capture discontinuities. An example of discontinuity analysis in PD treatment is reported by Strauss and his colleagues (Strauss et al., 2006). In their treatment study of patients with avoidant PD and obsessive-compulsive PD, they found that a worsening in the alliance followed by an improvement (a “rupture-repair” episode) predicted improvement in personality disorder symptoms and depression.

In the UPP, some of these constraints were met by collecting different kinds of data and having two follow-up evaluations, although all analyses used in the studies were based on linearity and the results reported were based on average change in the two treatment samples. In addition, we analysed subsamples to explore whether patients with more severe pathology showed a different pattern of change than the patients with less severe pathology at the 8 months follow-up. Unfortunately, the number of patients was too low to extend these subsample analysis, and three assessment points are not adequate for more thorough

analyses of discontinuity or individual fluctuations in change. The lack of such process measures and statistical procedures might be a limitation of the present studies, and should be subject for further exploration throughout the UPP period.

Summary of the methodological strengths and limitations

The validity study of the SIPP-118 had some limitations regarding lack of Norwegian normal population and samples with only axis I disorders. The comparison to the Dutch samples showed promising results for the Norwegian translations, but the validity of the instrument has to be reviewed further in larger and more diverse samples.

The efficacy studies contained an extensive range of assessments enable to measure differences and changes in the two treatments. The number of patients was acceptable, and the number of patients meeting to evaluations was considered high. Hence, the statistical power for detecting changes and differences between treatments was considered good. The studies lacked a comparison group from a no-treatment condition. This was considered both unethical and unrealistic, but it reduces the interpretation of the estimated change as solely due to the treatments in question. The treatments were considered quite similar to realistic treatment conditions for PD patients in Norway. Thus, the results may be generalizable to Norwegian settings, although this external validity might have decreased the statistical conclusion validity.

In examining the cost of treatment, the highly skewed distribution of the use of additional health services, making very broad confidence intervals, could have affected the possibility to find differences between treatments. In addition, the lack of differences in effect between treatments was a hindrance for cost-effectiveness analysis. On the other hand, the results give valuable information of taking additional costs into account when making clinical choices of treatment modality for patients with PDs.

5.2 Discussion of the results

Measuring core PD pathology

SIPP-118 is a long sought for assessment designed for measuring the core pathology of PDs. The structure and validity of the instrument showed promising results in the original paper by the designers (Verheul et al., 2008). The results from the study of the translated Norwegian version of SIPP-118 were comparable with the Dutch results. The PD group from Norway showed an almost similar profile as the Dutch PD group and differed

significantly from the Dutch normal population group on the facet level. The reliability was good on all five domains and most facets. Thus, with some adjustments the psychometric qualities seems promising for the Norwegian version of this instrument. On the other side, both the Dutch study and our study can be criticised for not making a distinction between axis I and axis II. Thus, the question concerning the instruments ability to differ between personality problems and general psychiatric disorder is still not answered satisfactory. The present study did not comprise a large enough sample to replicate the factor structure of the five higher order domains, and this should be a focus for further studies.

Additionally, one main focus when developing an assessment of personality pathology has been the ability to measure change. Normal personality has been viewed as stable and robust, and assessments have been focusing more on the stability of personality than on the ability to change. Introducing adaptive capacity as the core of PD pathology anticipates a heightened ability to capture change as the focus on heritability and generic personality traits is reduced. Paper I evaluated the validity of the SIPP-118, and further studies should evaluate the instruments ability to measure change.

Another factor of importance when evaluating a new instrument is the clinical utility. SIPP-118 is designed to be both an instrument used in research evaluating treatments, and to be of help in everyday clinical life working with PD patients. For clinical use, the fact that the instrument is built upon a theoretical foundation is considered a strength. The way SIPP-118 emphasises the strengths in the individuals adaptive capacity along with weaknesses, may also be useful when making individual plans and making treatment decisions. The ability to measure change, and to distinguish axis I from axis II is also considered important for the clinical utility. In addition, the questions of cut-off scores to distinguish between clinically significant pathology vs. normality, and if different SIPP-118 profiles can predict outcome of different treatments are still unexplored. More studies are needed to explore the potentials of SIPP-118 as a useful instrument. If SIPP-118 is to be applied as a screening tool, one sum score instead of five might increase its clinical utility. The Dutch research group has developed a short form of the SIPP-118 (SIPP-SF) for research purposes. This version might hold some of the qualities mentioned above, but the SIPP-SF has to our knowledge not yet been tested in clinical studies.

The effect of short-time treatment

At 8 months follow-up, change on clinical measures ranged from small to medium in both treatments (DHP: mean ES = .44 / OIP: mean ES = .36). Thus, the patient sample as a whole showed improvement over time, but there were no significant differences between treatments. According to our original hypotheses we expected that CP would be more beneficial even in the short run. This may have been an overly optimistic view as seen in the subsequent reflections, although the results from the 8 months follow-up were in the lower range compared to unpublished data from the Norwegian Network (N= 2248, Pedersen, personal communication). But even if the effect sizes were moderate, and there was no difference between treatments, there was a significant total change which is not to be underestimated. For comparison, the effect sizes for change measured by SCL-90-R in Bateman and Fonagy's study (Bateman & Fonagy, 1999) was ES = 0.18 at 6 months follow-up and ES = 0.51 at 12 months follow-up.

Of course, as already mentioned, one cannot be sure that the changes reported at the 8 months follow-up were due to specific effects of the two treatments, but the results were promising for the following evaluations. About one third of the total sample reported positive reliable changes on clinical measures at 8 months follow-up, whereas about one out of ten patients reported negative changes. These results were evident in both treatments. Even if there was a trend of positive change, the results also emphasise the importance of being aware of possible iatrogenic effects (Fonagy & Bateman, 2006). An important question was whether the patients who showed deterioration on clinical measures would continue this negative trend, or whether the increase in symptoms and interpersonal problems represents a temporary state on the way to a more positive development. In the study of Bateman and Fonagy (Bateman & Fonagy, 1999) the main effects in the partial hospitalisation group took place at 12 months follow-up and onwards. This supported a hypothesis of larger effects in the UPP at the 18 and 36 month follow-up.

The lack of difference in dropout rate between treatments at 8 months follow-up was counter to our hypothesis. The drop out rate in both groups was quite low (13.5% in DHP and 7.5% in OIP) compared to other studies (Leichsenring & Leibing, 2003a; Ogrodniczuk & Piper, 2001; Verheul & Herbrink, 2007). A multicentre study of day hospital treatment programs in Norway (Karterud et al., 2003) reported an average dropout rate of 24%. The drop out rate in DHP in our study was well below this level, and two of the premature

terminations were actually not dropouts in a traditional sense. These two patients terminated due to advice from the therapists since they were re-diagnosed as having Asperger's syndrome and organic frontal lobe syndrome (being wrongly assessed as avoidant PD). On the other hand, one might speculate if the drop-out rate was artificially low. The therapists were instructed to be more outreaching than their usual practice, which might have had an impact of the group culture, and thus may have had a negative impact on the treatment results.

Even more notable is the low dropout rate in the OIP treatment. Containment/alliance versus rupture/dropping out is a crucial factor in the treatment of patients with PD, and high dropout rates have been used as arguments against individual psychotherapy for patients with PDs (Giesen-Bloo et al., 2006; Linehan et al., 2006). Additionally, in intention to treat analysis dropouts tend to have a significant impact upon treatment results (Giesen-Bloo et al., 2006; Linehan et al., 2006). One possible explanation might be the flexible format of the OIP. Therapists were not forced to do any specific treatment, but were allowed to work according to their own preference and attuned to the most pressing needs of the patients. Of course, cautious must be taken when comparing dropout rate between studies due to different definitions of dropout, different procedures, and duration of treatment and follow-up periods (Shea et al., 1990), but this issue stands out as an important area for further exploration.

The effect of long-time treatment

The main finding in the 18 months follow-up evaluation was that the more intensive, multimodal combined psychotherapy (CP) was not more beneficial in terms of efficacy than outpatient individual psychotherapy (OIP). On the contrary, there was a trend towards better results in the OIP treatment. The results are consistent with one other study suggesting that a higher intensity of treatment dose might not yield significant additional effects, although this study compared inpatient treatment with a step-down model (Chiesa et al., 2004).

The termination rate was equal in the two treatments, and there was only one modest interaction effect between type of PD and type of treatment (in favour of OIP). The overall changes was in the moderate to high range (CP: mean ES = .51 with a range from ES = .27 - .92 / OIP: mean ES = .71 with a range from ES = .31 - 1.05). For comparison, Bateman and Fonagy (Bateman & Fonagy, 1999) found an effect size on SCL-90-R of ES = .56 after 18 months of mentalization based treatment for BPD patients (Bateman & Fonagy, 1999). In a

study of DBT versus community treatment by experts (CTBE) for BPD patients, changes on the Hamilton rating scale for depression were reported as $ES = .93$ (CTBE: $ES = .61$) at 12 months and $ES = .1.19$ at 24 months (CTBE: $ES = .88$) (Linehan et al., 2006). Patients in a CBT treatment study showed a change of $ES = .81$ on a symptom scale after 24 months of treatment (Davidson et al., 2006b; Davidson et al., 2006a), while Giesen-Bloo and colleagues reported $ES = 2.02$ on a mixed symptom scale at 3 years follow-up in a study of individual SFT for BPD patients (Giesen-Bloo et al., 2006). Although the overall effect sizes in our study was in the moderate to high range, and termination rate was comparable to other studies (Leichsenring & Leibing, 2003a; Ogrodniczuk & Piper, 2001; Verheul & Herbrink, 2007), the lack of difference between treatments was not in accordance with our hypotheses and was counterintuitive given the more intensive CP format. At 8 months follow-up we proposed various possible reasons for the lack of difference between treatments. One reason might be the transition phase after day hospital treatment, another might be that the intensive DHP stir up affects, and finally that a difference in change may be apparent in the long-term. One could argue that 18 months are still not long-term, but according to the reported development in treatment effect in other studies (e.g. Bateman & Fonagy, 2008b; Giesen-Bloo et al., 2006), a difference between treatments should be apparent after 1 ½ year of psychotherapy. The methodological limitations already discussed should of course be taken into consideration, but one might infer from these results that the decision of treatment modality for PD patients should be taken based on cost or availability considerations as the effect of the treatments is similar. But, before drawing such a conclusion, there are some important issues to address regarding the findings described.

First, all evidence based methods for PDs are highly structured, and there has been proposed that the coherence and structure of the therapy has more impact on efficacy than treatment theories. On the other hand, the structure of the CP treatment may have been negative for some patients. The initial extensive treatment requires that patients who are employed get a sick-leave for the treatment period, possibly withdrawing from important structures in their lives that could be important for their social functioning.

Furthermore, our study showed that almost 1/3 of the patients chose to end treatment before the 18 months follow-up independent of treatment allocation. This result raises a discussion of the benefit, or need, of long-term treatment per se, and illustrates the difficulty in maintaining treatment alliance in the work with PD patients. Many theories and treatment

models for patients with PDs have asserted long-term treatment as necessary for these patients. Long-term treatment has been supported by treatment studies, studies of the natural course of the disorder, as well as being implemented in the practice guidelines for BPD (American Psychiatric Association, 2001; Bateman & Fonagy, 2008b; Paris, 2002; Perry et al., 1999). On the other hand, short-term treatment for patients with PDs has also shown good results, even though these studies have been criticised for treating patients with less severe disorders that responds to short-term treatment, and especially cognitive treatment, in a better way than patients with more severe pathology (Abbass et al., 2008; Svartberg et al., 2004). Systematic knowledge about patient characteristics predicting the duration of long-term psychotherapy is largely absent, but Perry and colleagues (Perry et al., 2007) found in a recent study that dysthymic disorder, presence of any PD (particularly dependent PD), emotional neglect in childhood, and higher adaptive defence style predicted a greater number of sessions, while obsessive compulsive PD predicted fewer sessions. Even though our data showed that patients ending treatment before 18 months showed no more change than the patients still in treatment, one cannot simply assume that long-term treatment is needed for all patients with chronic disorders. Emerging questions might be how long is long-term treatment, and if short-term treatment may be adequate for some patients. These are empirical questions taking into account differences in severity of the disorder, type of diagnoses, and theoretical and practical development in the field.

Regarding the issue of alliance it can be assumed that the initial day hospital treatment increases the possibility for making a robust alliance as the day hospital treatment offers multiple arenas for containing the patients' interpersonal problems and early insecure attachment experiences. Increased number of therapists and patients are available for positive alliances, and teamwork among the therapists may reduce acting out countertransference feelings in the staff. Additionally, the patient's negative transference may be diluted. Also, the patients in CP experienced a break in therapeutic modalities from the initial 18 weeks treatment to the following treatment where the patients were confronted with the challenges of both separation and forming new attachment relationships. The OIP patients had the benefit of continuing their treatment with the same therapist, and not being influenced by complicated group dynamics. These other patients can also represent a source of support, comfort and identification. The attrition rate was not significantly different in the two treatments, but one may speculate if some of the patients in OIP may leave treatment and then resume treatment after a while. A treatment model, intermittent therapy, with more

tolerant boundaries regarding attrition has been suggested as an alternative model for patients with insecure attachment (Paris, 2007). Intermittent therapy is independent of any specific method of therapy, and involves low-frequency follow-up and support after a period of regular and consistent treatment (Paris, 2007). Intermittent therapy could be important in the way that therapists need to consider whether partial improvement can be considered “good enough”. There is no standard for mental health, and by giving the patients sufficient space for “acting out” and ending therapy, or by the therapist offering a break, patients might try out something in real life with the confidence that they can come back into therapy when needed. The discussion regarding degree of structured treatment versus the ability to individually monitor the patients’ progress needs further analysis.

In paper II the question regarding possible benefits for subgroups of patients were explored, but the size of the sample and the limited number of certain PDs prevented further exploration. We suggest that future studies also include exploration of individual change patterns, and nonlinear development.

The final topic to be raised is the quality of the treatments. For the follow-up conjoint individual and group treatment in CP the therapists have been interviewed about their experience of the conjoint treatment model, but these data has not yet been analysed. Anyhow, the general attitude seems to have been positive, and the therapists have expressed satisfaction with being part of a psychotherapy research project working within firm boundaries and structures, and not being stuck with difficult patients on their own. Typical difficulties have been tendencies to postpone meetings between the group and individual therapists, despite clear instructions, as well as difficulties in acknowledging the differences between one-modality treatment and conjoint treatment with regard to engaging patients in the dual work. The data obtained from this study suggest that the conjoint treatment model for this sample of PD patients does not seem superior to individual treatment alone, but the study should be replicated before any firm conclusions are drawn.

Health utilization costs

The main finding from this study as reported in paper IV was that treatment costs were higher, and additional costs were significantly lower in CP than in OIP during the first 8 months. The lowered additional costs in CP persisted at the 18 months follow-up resulting in equal total costs for the structured and extensive day hospital treatment followed by combined treatment as for the individual outpatient treatment when additional costs were

included in the calculations. Although the sample was small and highly skewed with the methodological and statistical limitations already discussed, these results are intriguing. Not only was the total costs for patients in the two treatments equal, but the total cost per month at 18 months follow-up were almost equal with the costs 12 months prior to treatment. If taking these calculations literally, it would mean that for this PD patient population extensive treatment is not more costly for the society than sporadic, unstructured treatment as was the pattern prior to the inclusion in the UPP.

In comparing alternative treatments for patients with PDs (as in our study), the approach that require less investment of time and resources would usually be preferred if the efficacy is similar. An alternative approach would be to see health service utilization costs as adding information to the results. This approach is grounded on an earlier study suggesting that day hospital treatment offer adequate containing capacity compared to inpatient treatment for patients with PDs (Karterud et al., 1992). They judged containing capacity by the frequency of dropouts, suicides, suicide attempts, frequency of psychotic breakdowns and level of medication. One could argue that health utilization costs is an indirect measure of containing capacity, and hypothesise that structured treatment, as CP, would offer a heightened containing capacity compared to a less structured treatment, as OIP. This might indicate a heightened containing capacity in the initial day hospital treatment.

The continuing decrease in use of additional services in CP is counter to the hypothesis of patient regression or patients forming an identity as hospitalised for these patients. In addition, the decrease might indicate a change in the patients which the clinical assessments were not able to capture. The methodological limitations make these reflections only suggestive, and the results certainly need to be replicated. The skewness in the use of health services indicates subgroups that use a lot more health services than other patients (20% of the patients used 85% of the total costs before inclusion in the project). If health utilization costs are a valid indicator of containing capacity, one might hypothesise that the need for containing capacity may be an indication of the choice of treatment modality for patients with PDs. The forthcoming evaluations of this sample of PD patients could answer some of these questions, and it will be interesting to follow the progress in the years to come. In addition, further studies are needed to explore these issues. Such studies should comprise large enough samples to perform subgroup analysis.

6. Conclusions and future directions

The cross-national validity study of the Norwegian version of the SIPP-118 showed adequate psychometric properties at the facet level and seems promising as a measure of PD core pathology. Future studies should include investigation of the instruments ability to distinguish between axis I and axis II pathology and its sensitivity to change.

The main findings in the two efficacy studies showed that at both 8 months follow-up and 18 months follow-up there were no differences between CP and OIP on clinical variables. This finding challenges the notion that extensive and structured day hospital treatment models followed by group and individual outpatient treatment are more efficient than outpatient individual treatment for patients with PDs. The limited analysis of sub groups showed no indication that the initial day hospital treatment was better for the most severely disturbed patients. Even though both internal and external validity was considered satisfactory regarding the efficacy studies, one should of course be cautious making health political inferences solely on this study. The studies should be replicated at different sites and in larger samples that make extended analyses of subgroups possible. The research should be supplemented with process studies exploring individual and non-linear changes, as well as predictors of change.

There was no difference in total costs for patients in CP and OIP at 18 months follow-up. The higher treatment costs in CP were compensated by a reduction in the use of additional health services. Future studies should comprise larger samples of patients and explore the differences between high cost patients and low cost patients. In addition, implementation of standardised measures makes it possible to compare costs across studies and diagnostic categories. Cost-benefit studies are to be preferred.

Even if there are obvious limitations in an extensive study such as the UPP, the results from this study underline the importance of carrying out such studies. The questions concerning different modalities for treatment for various degrees of severity and the different PD diagnoses, and the generalizability to clinical work at the psychiatric health centres favours a study design combining RCT design with a naturalistic approach. The upcoming 36 months evaluation of patient change and the process studies within the UPP will extend the presented results and bring knowledge into the field of day hospital treatment and general

PD treatment leading to a possibility for offering better treatment for the patients in need of help.

7. References

- Abbass, A., Sheldon, A., Gyra, J., & Kalpin, A. (2008). Intensive short-term dynamic psychotherapy for DSM-IV personality disorders: a randomized controlled trial. *Journal of Nervous & Mental Disease*, 196, 211-216.
- Alden, L. E., Wiggins, J. S., & Pincus, A. L. (1990). Construction of Circumplex scales for the Inventory of Interpersonal Problems. *Journal of Personality Assessment*, 55, 521-536.
- American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental disorders*. (4 ed.) Washington, DC: American Psychiatric Press.
- American Psychiatric Association (2001). Practice guideline for the treatment of patients with borderline personality disorder. American Psychiatric Association. *American Journal of Psychiatry*, 158, Suppl-52.
- Andrea, H., Verheul, R., Berghout, C., Dolan, C., Vanderkroft, P., Busschbach, J. J. et al. (2007). *Measuring the Core components of Maladaptive Personality: Severity Indices of Personality Problems (SIPP-118)* <http://hdl.handle.net/1765/10066>: Report of the Viersprong Institute for studies on Personality Disorders (VISPD) in cooperation with the Department of Medical Psychology & Psychotherapy, Erasmus University Rotterdam, The Netherlands.
- Bartak, A., Soeteman, D. I., Verheul, R., & Busschbach, J. J. (2007). Strengthening the status of psychotherapy for personality disorders: an integrated perspective on effects and costs. *Canadian Journal of Psychiatry - Revue Canadienne de Psychiatrie*, 52, 803-810.
- Bateman, A. & Fonagy, P. (1999). Effectiveness of partial hospitalization in the treatment of borderline personality disorder: a randomized controlled trial. *American Journal of Psychiatry*, 156, 1563-1569.
- Bateman, A. W. & Fonagy, P. (2000). Effectiveness of psychotherapeutic treatment of personality disorder. *British Journal of Psychiatry*, 177, 138-143.
- Bateman, A. & Fonagy, P. (2001). Treatment of borderline personality disorder with psychoanalytically oriented partial hospitalization: an 18-month follow-up. *American Journal of Psychiatry*, 158, 36-42.
- Bateman, A. (2003). Health service utilization costs for borderline personality disorder patients treated with psychoanalytically oriented partial hospitalization versus general psychiatric care. *The American journal of psychiatry*, 160, 169-171.
- Bateman, A. & Fonagy, P. (2008a). 8-Year Follow-Up of Patients Treated for Borderline Personality Disorder: Mentalization-Based Treatment Versus Treatment as Usual. *American Journal of Psychiatry*.
- Bateman, A. & Fonagy, P. (2008b). Comorbid antisocial and borderline personality disorders: mentalization-based treatment. *Journal of Clinical Psychology*, 64, 181-194.
- Beck, A. T., Rush, A. J., & Shaw, B. F. (1979). *Cognitive therapy of depression*. New York: Guilford.

- Beck, A. T. & Steer, R. A. (1988). *Manual for the Beck Hopelessness Scale*. San Antonio Texas: Psychological Corporation.
- Beck, A., Ward, C., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Arch Gen Psychiatry*, 4, 561-571.
- Beecham, J., Sled, M., Knapp, M., Chiesa, M., & Drahorad, C. (2006). The costs and effectiveness of two psychosocial treatment programmes for personality disorder: a controlled study. *European psychiatry*, 21, 102-109.
- Bender, D. S., Dolan, R. T., Skodol, A. E., Sanislow, C. A., Dyck, I. R., McGlashan, T. H. et al. (2001). Treatment utilization by patients with personality disorders. *American Journal of Psychiatry*, 158, 295-302.
- Bender, D. S., Skodol, A. E., Pagano, M. E., Dyck, I. R., Grilo, C. M., Shea, M. T. et al. (2006). Prospective assessment of treatment use by patients with personality disorders. *Psychiatric Services*, 57, 254-257.
- Benjamin, L. S., Rothweiler, J. C., & Critchfield, K. L. (2006). The use of structural analysis of social behavior (SASB) as an assessment tool. [Review] [94 refs]. *Annual Review of Clinical Psychology*, 2, 83-109.
- Blum, N., St, J. D., Pfohl, B., Stuart, S., McCormick, B., Allen, J. et al. (2008). Systems Training for Emotional Predictability and Problem Solving (STEPPS) for outpatients with borderline personality disorder: a randomized controlled trial and 1-year follow-up.[see comment]. *American Journal of Psychiatry*, 165, 468-478.
- Brazier, J., Tumor, I., Holmes, M., Ferriter, M., Parry, G., Dent-Brown, K. et al. (2006). Psychological therapies including dialectic behaviour therapy for borderline personality disorder: a systematic review and preliminary economic evaluation. *Health Technology Assessment*, 10.
- Cantoni, E. & Ronchetti, E. (2004). *A robust approach for skewed and heavy-tailed outcomes in the analysis of health care expenditures* (Rep. No. 2004.03). Département d'économétrie: Université de Genève.
- Chiesa, M., Fonagy, P., Holmes, J., Drahorad, C., & Harrison-Hall, A. (2002). Health service use costs by personality disorder following specialist and nonspecialist treatment: a comparative study. *Journal of Personality Disorders*, 16, 160-173.
- Chiesa, M., Fonagy, P., Holmes, J., & Drahorad, C. (2004). Residential versus community treatment of personality disorders: a comparative study of three treatment programs. *American Journal of Psychiatry*, 161, 1463-1470.
- Clark, L. A., Watson, D., & Reynolds, S. (1995). Diagnosis and classification of psychopathology: challenges to the current system and future directions. *Annual Review of Psychology*, 46, 121-153.
- Clark, D. M. (1999). Anxiety disorders: why they persist and how to treat them. *Behaviour research and therapy*, 37, 5-27.

- Clarkin, J. F., Levy, K. N., Lenzenweger, M. F., & Kernberg, O. F. (2004). The Personality Disorders Institute/Borderline Personality Disorder Research Foundation randomized control trial for borderline personality disorder: rationale, methods, and patient characteristics. *Journal of Personality Disorders*, 18, 52-72.
- Clarkin, J. F., Levy, K. N., Lenzenweger, M. F., & Kernberg, O. F. (2007). Evaluating Three Treatments for Borderline Personality Disorder: A Multiwave Study. *American Journal of Psychiatry*, 164, 922-928.
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. (second ed.) Hillsdale, NJ: Lawrence Earlbaum Associates.
- Coid, J. (2003). Epidemiology, public health and the problem of personality disorder. *Br.J.Psychiatry Suppl*, 44, S3-10.
- Coid, J., Yang, M., Tyrer, P., Roberts, A., & Ullrich, S. (2006). Prevalence and correlates of personality disorder in Great Britain. *British Journal of Psychiatry*, 188, 423-431.
- Davidson, K., Norrie, J., Tyrer, P., Gumley, A., Tata, P., Murray, H. et al. (2006a). The effectiveness of cognitive behavior therapy for borderline personality disorder: results from the borderline personality disorder study of cognitive therapy (BOSCOT) trial. *Journal of Personality Disorders*, 20, 450-465.
- Davidson, K., Tyrer, P., Gumley, A., Tata, P., Norrie, J., Palmer, S. et al. (2006b). A randomized controlled trial of cognitive behavior therapy for borderline personality disorder: rationale for trial, method, and description of sample. *Journal of Personality Disorders*, 20, 431-449.
- Dawson, D. A., Grant, B. F., Stinson, F. S., & Chou, P. S. (2005). Psychopathology associated with drinking and alcohol use disorders in the college and general adult populations. *Drug & Alcohol Dependence*, 77, 139-150.
- Derogatis, L. R. (1983). *SCL-90-R Manual: Administration, Scoring & procedures*. USA: Clinical psychometric research.
- Dolan-Sewell, R. T., Krueger, R. F., & Shea, M. T. (2001). Co-occurrence with syndrome disorders. In W.J.Livesley (Ed.), *Handbook of personality disorders : theory, research, and treatment* (pp. 84-104). New York: Guilford Press.
- Ekselius, L., Tillfors, M., Furmark, T., & Fredrikson, M. (2001). Personality disorders in the general population: DSM-IV and ICD-10 defined prevalence as related to sociodemographic profile. *Personality and Individual Differences*, 311-320.
- Emmelkamp, P. M., Benner, A., Kuipers, A., Feiertag, G. A., Koster, H. C., & van Apeldoorn, F. J. (2006). Comparison of brief dynamic and cognitive-behavioural therapies in avoidant personality disorder. *Br J Psychiatry*, 189, 60-64.
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioural, and biomedical sciences. *Behavioural Research Methods*, 39, 175-191.

First, M. B. (1994). *Structural clinical interview for DSM-IV Axis II personality disorders (version 2.0)* New York: New York State Psychiatric Institute.

Fonagy, P., Target, M., Gergely, G., & Jurist, E. J. (2002). *Affect Regulation, Mentalization and the Development of the Self*. New York: Other Press.

Fonagy, P. & Bateman, A. (2006). Progress in the treatment of borderline personality disorder. *Br J Psychiatry*, 188, 1-3.

Frankenburg, F. R. & Zanarini, M. C. (2004). The association between borderline personality disorder and chronic medical illnesses, poor health-related lifestyle choices, and costly forms of health care utilization. *Journal of Clinical Psychiatry*, 65, 1660-1665.

Giesen-Bloo, J., van Dyck, R., Spinhoven, P., van Tilburg, W., Dirksen, C., van Asselt, T. et al. (2006). Outpatient Psychotherapy for Borderline Personality Disorder: Randomized Trial of Schema-Focused Therapy vs Transference-Focused Psychotherapy. *Archives of General Psychiatry*, 63, 649-658.

Goldman, H. H., Skodol, A. E., & Lave, T. R. (1992). Revising Axis V for DSM-IV: A review of measures of social functioning. *American Journal of Psychiatry*, 149, 1148-1156.

Grant, B. F., Hasin, D. S., Stinson, F. S., Dawson, D. A., Chou, S. P., Ruan, W. J. et al. (2004). Prevalence, correlates, and disability of personality disorders in the United States: results from the national epidemiologic survey on alcohol and related conditions. *Journal of Clinical Psychiatry*, 65, 948-958.

Grant, B. F., Stinson, F. S., Dawson, D. A., Chou, S. P., & Ruan, W. J. (2005). Co-occurrence of DSM-IV personality disorders in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Comprehensive Psychiatry*, 46, 1-5.

Gråwe, R. W., Hagen, H., Husum, T. L., Pedersen, P. B., & Ruud, T. (2005). *Pasienter som mottar poliklinisk behandling i psykisk helsevern for voksne - September 2004* (Rep. No. STF778 A055013). http://www.sintef.no/content/page1_6304.aspx: SINTEF Helse.

Groth-Marnat, G. (1990). *The Handbook of Psychological Assessment*. (Second ed.) New York: John Wiley & Sons.

Halsteinli, V., Hatling, T., & Kalleth, J. (2003). *Aktivitetsbaserte poliklinikkter i psykisk helsevern - mellom visjon og virkelighet* (Rep. No. STF78 A035001). SINTEF Unimed.

Hayes, A. M., Laurenceau, J. P., Feldman, G., Strauss, J. L., & Cardaciotto, L. (2007). Change is not always linear: the study of nonlinear and discontinuous patterns of change in psychotherapy. *Clinical Psychology Review*, 27, 715-723.

Hoffart, A., Hauge, T., Hedley, L., & Larsen, S. (1993). Kognitiv terapi ved panikk med agorafobi. *Tidsskrift for Norsk Psykologforening*, 30, 742-748.

Holden, R. R. & Fekken, G. C. (1988). Test-retest reliability of the Hopelessness Scale and its items in a university population. *J Clin. Psychol.*, 44, 40-43.

- Horowitz, L. M., Rosenberg, S. E., Baer, B. A., Ureno, G., & Villaseñor, V. S. (1988). Inventory of interpersonal problems: psychometric properties and clinical applications. *J.Consult Clin.Psychol.*, 56, 885-892.
- Hudson, W. W. (1982). *The clinical measurement package: A field manual*. Chicago: The Dorsey Press.
- Hummelen, B., Wilberg, T., & Karterud, S. (2007). Interviews of female patients with borderline personality disorder who dropped out of group psychotherapy. *Int.J Group Psychother.*, 57, 67-91.
- Husum, T. L., Gråwe, R. W., & Bjørngaard, J. H. (2005). *Avtalespesialister i det psykiske helsevernet i Midt-Norge: Kartlegging av helsetilbudet, samhandling og tilgjengelighet* (Rep. No. STF78 A045022). http://www.sintef.no/content/page1_4620.aspx: SINTEF Helse.
- Jackson, H. J. & Burgess, P. M. (2000). Personality disorders in the community: a report from the Australian National Survey of Mental Health and Wellbeing. *Social Psychiatry and Psychiatric Epidemiology*, 35, 531-538.
- Jacobson, N. S. & Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin.Psychol.*, 59, 12-19.
- Johns, S. & Karterud, S. (2004). Guidelines for Art Group Therapy as Part of a Day Treatment Program for Patients with Personality Disorders. *Group Analysis*, 37, 419-432.
- Kallert, T. W., Glockner, M., Priebe, S., Briscoe, J., Rymaszewska, J., Adamowski, T. et al. (2004a). A comparison of psychiatric day hospitals in five European countries: implications of their diversity for day hospital research. *Social Psychiatry and Psychiatric Epidemiology*, 39, 777-788.
- Kallert, T. W., Matthes, C., Glockner, M., Eichler, T., Koch, R., & Schutzwahl, M. (2004b). [Acute psychiatric day hospital treatment: is the effectiveness of this treatment approach still questionable?]. *Psychiatrische Praxis*, 31, 409-419.
- Karterud, S., Vaglum, S., Friis, S., Irion, T., Johns, S., & Vaglum, P. (1992). Day hospital therapeutic community treatment for patients with personality disorders. An empirical evaluation of the containment function. *J Nerv.Ment.Dis.*, 180, 238-243.
- Karterud, S., Pedersen, G., Bjordal, E., Brabrand, J., Friis, S., Haaseth, O. et al. (2003). Day treatment of patients with personality disorders: experiences from a Norwegian treatment research network. *J Personal.Disord.*, 17, 243-262.
- Karterud, S., Johansen, M. S., & Wilberg, T. (2007). Conjoint group and individual psychotherapy in a research trial for patients with severe personality disorders. *Group*, 31, 31-46.
- Kernberg, O. (1984). *Severe personality disorders*. New haven, CT: Yale University Press.
- Knekt, P., Lindfors, O., Laaksonen, M. A., Raitasalo, R., Haaramo, P., Jarvikoski, A. et al. (2008). Effectiveness of short-term and long-term psychotherapy on work ability and

functional capacity--a randomized clinical trial on depressive and anxiety disorders. *Journal of Affective Disorders*, 107, 95-106.

Koerner, K. & Linehan, M. M. (2000). Research on dialectical behavior therapy for patients with borderline personality disorder. *Psychiatric Clinics of North America*, 23, 151-167.

Krawitz, R. (1997). A prospective psychotherapy outcome study. *Aust.N.Z.J Psychiatry*, 31, 465-473.

Krueger, R. F. (2005). Continuity of axes I and II: toward a unified model of personality, personality disorders, and clinical disorders. *Journal of Personality Disorders*, 19, 233-261.

Leichsenring, F. & Leibling, E. (2003). The effectiveness of psychodynamic therapy and cognitive behavior therapy in the treatment of personality disorders: a meta-analysis. *American Journal of Psychiatry*, 160, 1223-1232.

Lenzenweger, M. F. & Clarkin, J. F. (2005). *Major Theories of Personality Disorder*. (2 ed.) New York: The Guilford Press.

Linehan, M. M., Armstrong, H. E., Suarez, A., Allmon, D., & Heard, H. L. (1991). Cognitive-behavioral treatment of chronically parasuicidal borderline patients. *Archives of General Psychiatry*, 48, 1060-1064.

Linehan, M. M., Heard, H. L., & Armstrong, H. E. (1993). Naturalistic follow-up of a behavioral treatment for chronically parasuicidal borderline patients. *Arch Gen Psychiatry*, 50, 971-974.

Linehan, M. M., Tutek, D. A., Heard, H. L., & Armstrong, H. E. (1994). Interpersonal outcome of cognitive behavioral treatment for chronically suicidal borderline patients. *Am J Psychiatry*, 151, 1771-1776.

Linehan, M. M., Comtois, K. A., Murray, A. M., Brown, M. Z., Gallop, R. J., Heard, H. L. et al. (2006). Two-year randomized controlled trial and follow-up of dialectical behavior therapy vs therapy by experts for suicidal behaviors and borderline personality disorder. *Arch Gen Psychiatry*, 63, 757-766.

Livesley, W. J., Jang, K. L., & Vernon, P. A. (1998). Phenotypic and genetic structure of traits delineating personality disorder. *Archives of General Psychiatry*, 55, 941-948.

Livesley, W. J. & Jang, K. L. (2000). Toward an empirically based classification of personality disorder. *Journal of Personality Disorders*, 14, 137-151.

Livesley, W. J. (2001). Commentary on reconceptualizing personality disorder categories using trait dimensions. *Journal of Personality*, 69, 277-286.

Livesley, W. J. & Jang, K. L. (2005). Differating normal, abnormal and disordered personality. *European journal of personality*, 19, 257-268.

Lynch, T. R., Chapman, A. L., Rosenthal, M. Z., Kuo, J. R., & Linehan, M. M. (2006). Mechanisms of change in dialectical behavior therapy: Theoretical and empirical observations. *Journal of Clinical Psychology*, 62, 459-480.

- Lynch, T. R., Trost, W. T., Salsman, N., & Linehan, M. M. (2007). Dialectical behavior therapy for borderline personality disorder. *Annual Review of Clinical Psychology*, rev., clin-205.
- Marks, I. M. (1987). *Fears phobias and rituals*. Oxford: Oxford University Press.
- Marshall, M., Crowther, R., Almaraz-Serrano, A. M., & Tyrer, P. (2001). Day hospital versus out-patient care for psychiatric disorders. *The Cochrane database of systematic reviews*, CD003240.
- McCrae, R. R., Costa, P. T., Jr., Ostendorf, F., Angleitner, A., Hrebickova, M., Avia, M. D. et al. (2000). Nature over nurture: temperament, personality, and life span development. *Journal of Personality & Social Psychology*, 78, 173-186.
- McCrae, R. R., Yang, J., Costa, P. T., Jr., Dai, X., Yao, S., Cai, T. et al. (2001). Personality profiles and the prediction of categorical personality disorders. *Journal of Personality*, 69, 155-174.
- McGlashan, T. H., Grilo, C. M., Sanislow, C. A., Ralevski, E., Morey, L. C., Gunderson, J. G. et al. (2005). Two-year prevalence and stability of individual DSM-IV criteria for schizotypal, borderline, avoidant, and obsessive-compulsive personality disorders: toward a hybrid model of axis II disorders. *American Journal of Psychiatry*, 162, 883-889.
- Melberg, H. O., Lauritzen, L., & Ravndal, E. (2003). *Hvilken nytte, for hvem og til hvilken kostnad? En prospektiv studie av stoffmisbrukere i behandling* (Rep. No. 4). Oslo: SIRUS.
- Moffett, L. A. & Radenhausen, R. A. (1990). Assessing depression in substance abusers: Beck Depression Inventory and SCL-90R. *Addictive Behaviors*, 15, 179-181.
- Mullins-Sweatt, S. N. & Widiger, T. A. (2008). Comparative validity of the Shedler and Westen Assessment Procedure-200. *Psychological assessment*, 20, 183-188.
- Narud, K., Mykletun, A., & Dahl, A. A. (2005). A comparison of gains after treatment at a psychiatric outpatient clinic in patients with cluster A + B, or cluster C personality disorders, and non-psychotic axis I disorders. *Nordic Journal of Psychiatry*, 59, 331-338.
- Nestadt, G., Costa, P. T., Jr., Hsu, F. C., Samuels, J., Bienvenu, O. J., & Eaton, W. W. (2008). The relationship between the five-factor model and latent Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition personality disorder dimensions. *Comprehensive Psychiatry*, 49, 98-105.
- NLWA (2008). *Hva foregår på legekantorene? Konsultasjonsstatistikk for 2006*. (Rep. No. 4). The Norwegian Labour and Welfare Administration.
- Ogrodniczuk, J. S. & Piper, W. E. (2001). Day treatment for personality disorders: a review of research findings. *Harvard Review of Psychiatry*, 9, 105-117.
- Oldham, J. M., Skodol, A. E., Kellman, H. D., Hyler, S. E., Doidge, N., Rosnick, L. et al. (1995). Comorbidity of axis I and axis II disorders. *American Journal of Psychiatry*, 152, 571-578.

Paris, J. (2002). Implications of long-term outcome research for the management of patients with borderline personality disorder. *Harvard Review of Psychiatry*, 10, 315-323.

Paris, J. (2007). Intermittent psychotherapy: an alternative to continuous long-term treatment for patients with personality disorders. *J.Psychiatr.Pract.*, 13, 153-158.

Pedersen, G. (2000). Circumplex of Interpersonal Problems (CIP): Norsk revidert versjon av Inventory of Interpersonal Problems-Circumplex (IIP-C). *Journal of Norwegian Psychological Association*, 39, 25-34.

Pedersen, G. (2007a). Test-retest Quality of life scale, unpublished data. 2007a. Ref Type: Personal Communication

Pedersen, G., Hagtvet, K. A., & Karterud, S. (2007a). Generalizability studies of the Global Assessment of Functioning-Split version. *Comprehensive Psychiatry*, 48, 88-94.

Pedersen, G., Hagtvet, K. A., & Karterud, S. (2007b). Self-therapist concordance and therapist agreement on patients interpersonal problems. In *The 15th clinical conference for the Norwegian Network of Psychotherapy Day Hospitals*.

Pedersen, P. B. (2007b). *SAMDATA Sektorrapport for det psykiske helsevernet 2006* (Rep. No. 2/07). SINTEF Health Research.

Perry, J. C. (1993). Longitudinal studies of personality disorders. *Journal of Personality Disorders*, 7, 63-85.

Perry, J. C., Banon, E., & Ianni, F. (1999). Effectiveness of psychotherapy for personality disorders. *American Journal of Psychiatry*, 156, 1312-1321.

Perry, J. C., Bond, M., & Roy, C. (2007). Predictors of treatment duration and retention in a study of long-term dynamic psychotherapy: childhood adversity, adult personality, and diagnosis. *J.Psychiatr.Pract.*, 13, 221-232.

Petersen, S. Ø. (2007). *SAMDATA Sektorrapport for somatisk spesialisthelsetjeneste 2006* (Rep. No. 3/07). SINTEF Health Research.

Piper, W. E., Rosie, J. S., Azim, H. F., & Joyce, A. S. (1993). A randomized trial of psychiatric day treatment for patients with affective and personality disorders. *Hospital and Community Psychiatry*, 44, 757-763.

Rendu, A., Moran, P., Patel, A., Knapp, M., & Mann, A. (2002). Economic impact of personality disorders in UK primary care attenders. *British Journal of Psychiatry*, 181, 62-66.

Robins, C. J. & Chapman, A. L. (2004). Dialectical behavior therapy: current status, recent developments, and future directions. *Journal of Personality Disorders*, 18, 73-89.

Rossberg, J. I., Karterud, S., Pedersen, G., & Friis, S. (2007). An empirical study of countertransference reactions toward patients with personality disorders. *Comprehensive Psychiatry*, 48, 225-230.

- Seligman, M. E. P. (1995). The effectiveness of psychotherapy: The consumer reports study. *American Psychologist*, 50, 965-974.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin Company.
- Shea, M. T., Pilkonis, P. A., Beckham, E., Collins, J. F., Elkin, I., Sotsky, S. M. et al. (1990). Personality disorders and treatment outcome in the NIMH Treatment of Depression Collaborative Research Program. *Am J Psychiatry*, 147, 711-718.
- Shea, M. T., Stout, R., Gunderson, J., Morey, L. C., Grilo, C. M., McGlashan, T. et al. (2002). Short-term diagnostic stability of schizotypal, borderline, avoidant, and obsessive-compulsive personality disorders. *American Journal of Psychiatry*, 159, 2036-2041.
- Sheehan, D. V., Lecrubier, Y., Janavs, J., Knapp, E., Weiller, E., & Bonora, L. I. (1994). *Mini International Neuropsychiatric Interview (M.I.N.I)* Tampa, Florida and Paris, France.
- Skodol, A. E., Gunderson, J. G., McGlashan, T. H., Dyck, I. R., Stout, R. L., Bender, D. S. et al. (2002a). Functional impairment in patients with schizotypal, borderline, avoidant, or obsessive-compulsive personality disorder. *American Journal of Psychiatry*, 159, 276-283.
- Skodol, A. E., Siever, L. J., Livesley, W. J., Gunderson, J. G., Pfohl, B., & Widiger, T. A. (2002b). The borderline diagnosis II: biology, genetics, and clinical course. *Biological Psychiatry*, 51, 951-963.
- Skodol, A. E., Oldham, J. M., Bender, D. S., Dyck, I. R., Stout, R. L., Morey, L. C. et al. (2005ab). Dimensional Representations of DSM-IV Personality Disorders: Relationships to Functional Impairment. *American Journal of Psychiatry*, 162, 1919-1925.
- Skodol, A. E., Pagano, M. E., Bender, D. S., Shea, M. T., Gunderson, J. G., Yen, S. et al. (2005c). Stability of functional impairment in patients with schizotypal, borderline, avoidant, or obsessive-compulsive personality disorder over two years. *Psychological Medicine*, 35, 443-451.
- Smith, T. L. & Benjamin, L. S. (2002). The functional impairment associated with personality disorders. *Current Opinion in Psychiatry*, 15, 135-141.
- Soeteman, D. I., Timman, R., Trijsburg, R. W., Verheul, R., & Busschbach, J. J. (2005). Assessment of the burden of disease among inpatients in specialized units that provide psychotherapy. *Psychiatric Services*, 56, 1153-1155.
- Stevenson, J. (1999). Psychotherapy with borderline patients: II. A preliminary cost benefit study. *Australian and New Zealand Journal of Psychiatry*, 33, 473-477.
- Stolk, E. A., Brouwer, W. B., & Busschbach, J. J. (2002). Rationalising rationing: economic and other considerations in the debate about funding of Viagra. *Health Policy*, 59, 53-63.
- Strauss, J. L., Hayes, A. M., Johnson, S. L., Newman, C. F., Brown, G. K., Barber, J. P. et al. (2006). Early alliance, alliance ruptures, and symptom change in a nonrandomized trial of cognitive therapy for avoidant and obsessive-compulsive personality disorders. *J.Consult Clin.Psychol.*, 74, 337-345.

Svartberg, M., Stiles, T. C., & Seltzer, M. H. (2004). Randomized, controlled trial of the effectiveness of short-term dynamic psychotherapy and cognitive therapy for cluster C personality disorders. *American Journal of Psychiatry*, 161, 810-817.

Svrakic, D. M., Draganic, S., Hill, K., Bayon, C., Przybeck, T. R., & Cloninger, C. R. (2002). Temperament, character, and personality disorders: etiologic, diagnostic, treatment issues. *Acta Psychiatrica Scandinavica*, 106, 189-195.

Torgersen, S. (1995). *Personlighet og personlighetsforstyrrelser*. Oslo: Universitetsforlaget.

Torgersen, S., Kringlen, E., & Cramer, V. (2001). The prevalence of personality disorders in a community sample. *Archives of General Psychiatry*, 58, 590-596.

Trull, T. J. (2005). Dimensional models of personality disorder: coverage and cutoffs. *Journal of Personality Disorders*, 19, 262-282.

Tyrer, P., Tom, B., Byford, S., Schmidt, U., Jones, V., Davidson, K. et al. (2004). Differential effects of manual assisted cognitive behavior therapy in the treatment of recurrent deliberate self-harm and personality disturbance: the POPMACT study. *Journal of Personality Disorders*, 18, 102-116.

Tyrer, P. (2005). The problem of severity in the classification of personality disorder. *Journal of Personality Disorders*, 19, 309-314.

van Asselt, A. D. I., Dirksen, C. D., Arntz, A., & Severens, J. L. (2007). The cost of borderline personality disorder: Societal cost of illness in BPD-patients. *European psychiatry*, 22, 354-361.

Verheul, R., Van Den Bosch, L. M., Koeter, M. W., De Ridder, M. A., Stijnen, T., & Van Den, B. W. (2003). Dialectical behaviour therapy for women with borderline personality disorder: 12-month, randomised clinical trial in The Netherlands. *Br J Psychiatry*, 182, 135-140.

Verheul, R. (2005). Clinical utility of dimensional models for personality pathology. *Journal of Personality Disorders*, 19, 283-302.

Verheul, R. (2006). Changeability of personality disorders.
Ref Type: Personal Communication

Verheul, R. & Herbrink, M. (2007). The efficacy of various modalities of psychotherapy for personality disorders: a systematic review of the evidence and clinical recommendations. *Int.Rev.Psychiatry*, 19, 25-38.

Verheul, R., Andrea, H., Berghout, C. C., Dolan, C., Busschbach, J. J., van der Kroft, P. J. et al. (2008). Severity Indices of Personality Problems (SIPP-118): development, factor structure, reliability, and validity. *Psychol.Assess.*, 20, 23-34.

Walmyr Publishing CO (1992). *The Walmyr Assessment Scale Scoring Manual* Tempe, AZ.

Walters, S. J. & Campbell, M. J. (2004). The use of bootstrap methods for analysing Health-Related Quality of Life outcomes (particularly the SF-36). *Health Qual.Life Outcomes.*, 2, 70.

- Westen, D. & Shedler, J. (2007). Personality diagnosis with the Shedler-Westen Assessment Procedure (SWAP): integrating clinical and statistical measurement and prediction. *Journal of Abnormal Psychology, 116*, 810-822.
- Widiger, T. A. & Costa, P. T., Jr. (1994). Personality and personality disorders. *Journal of Abnormal Psychology, 103*, 78-91.
- Widiger, T. A. & Simonsen, E. (2005). Alternative dimensional models of personality disorder: finding a common ground. *Journal of Personality Disorders, 19*, 110-130.
- Widiger, T. A. & Lowe, J. R. (2007). Five-factor model assessment of personality disorder. [Review] [102 refs]. *Journal of Personality Assessment, 89*, 16-29.
- Wilberg, T., Friis, S., Karterud, S., Mehlum, L., Urnes, O., & Vaglum, P. (1998). Patterns of short-term course in patients treated in a day unit for personality disorders. *Comprehensive Psychiatry, 39*, 75-84.
- Wilberg, T., Urnes, O., Friis, S., Irion, T., Pedersen, G., & Karterud, S. (1999). One-year follow-up of day treatment for poorly functioning patients with personality disorders. *Psychiatric Services, 50*, 1326-1330.
- Wilberg, T., Karterud, S., Pedersen, G., Urnes, O., Irion, T., Brabrand, J. et al. (2003). Outpatient group psychotherapy following day treatment for patients with personality disorders. *J Personal. Disord., 17*, 510-521.
- Young, J. (1999). *Cognitive therapy for Personality Disorders*. Sarasota, FL: Professional Resources Press.
- Zanarini, M. C. & Frankenburg, F. R. (2001). Attainment and maintenance of reliability of axis I and II disorders over the course of a longitudinal study. *Comprehensive Psychiatry, 42*, 369-374.
- Zanarini, M. C., Frankenburg, F. R., Hennen, J., & Silk, K. R. (2003). The longitudinal course of borderline psychopathology: 6-year prospective follow-up of the phenomenology of borderline personality disorder. *American Journal of Psychiatry, 160*, 274-283.
- Zeech, A., Hartmann, A., & Kuhn, K. (2005). Psychotherapy in a day clinic: Results of a 1.5 year follow-up. *Psychiatric Quarterly, 76*, 1-17.
- Zimmerman, M., Rothschild, L., & Chelminski, I. (2005). The Prevalence of DSM-IV Personality Disorders in Psychiatric Outpatients. *American Journal of Psychiatry, 162*, 1911-1918.

